

Université Paris 1 Panthéon-Sorbonne
École doctorale de Philosophie



THÈSE

pour l'obtention du grade de docteur en Philosophie
de l'Université de Paris 1 Panthéon-Sorbonne

Présentée et soutenue publiquement par

Marion VORMS

le 5 décembre 2009

THÉORIES, MODES D'EMPLOI

**UNE PERSPECTIVE COGNITIVE SUR L'ACTIVITÉ THÉORIQUE
DANS LES SCIENCES EMPIRIQUES**

Sous la co-direction de **Jean GAYON** et d'**Anouk BARBEROUSSE**

Composition du jury :

Anouk BARBEROUSSE	Chargée de recherche au CNRS, IHPST
Olivier DARRIGOL	Directeur de recherche au CNRS, REHSEIS (rapporteur)
Jean GAYON	Professeur à l'Université Paris 1 Panthéon-Sorbonne, IHPST
Paul W. HUMPHREYS	Professeur à l'Université de Virginie, Charlottesville
Pierre JACOB	Directeur de recherche au CNRS, Institut Jean Nicod (président du jury)

Année : 2009

Marion VORMS

THÉORIES, MODES D'EMPLOI

UNE PERSPECTIVE COGNITIVE SUR L'ACTIVITÉ THÉORIQUE
DANS LES SCIENCES EMPIRIQUES

THÈSE

pour l'obtention du grade de docteur en Philosophie
de l'Université de Paris 1 Panthéon-Sorbonne

Sous la co-direction
de Jean GAYON et d'Anouk BARBEROUSSE

*À la mémoire de mes grands-parents,
Matla Liba Oerbach, Hersz Ber Rapoport,
Andrée Van Praag et Jean Vorms.*

Table des matières

Remerciements	xi
Introduction générale : qu'est-ce qu'une théorie scientifique ?	1
1 Présupposés adoptés	1
1.1 Les théories sont des outils de représentation et d'inférence . . .	1
1.2 Représentation, explication et compréhension	3
2 Thèses défendues	6
2.1 La notion de théorie ne permet pas de rendre compte de l'ac- tivité scientifique dans la perspective des agents	6
2.2 Une hypothèse théorique : la notion de version	8
2.3 Un outil d'analyse : la notion de format	9
3 État de la question et méthodologie	10
3.1 L'héritage kuhnien	11
3.2 Les philosophes de la science « en pratique »	11
3.3 Les approches cognitives de l'activité scientifique	13
3.4 Les études sur la science (<i>science studies</i>)	13
4 Plan de l'argumentation	15
4.1 Première partie – Formulations et versions de la mécanique classique	15
4.2 Deuxième partie – Unités et outils d'analyse de l'activité théo- rique	16
4.3 Troisième partie – Représentations en génétique classique . . .	17
I Formulations et versions de la mécanique classique	19
Vue d'ensemble de la première partie	21
1 Quel est le contenu de la mécanique classique ?	23
1 Identifier le contenu de la mécanique classique	24

2	Première difficulté : les différentes formulations contemporaines de la mécanique classique	25
3	Deuxième difficulté : les présentations successives de la mécanique au cours de l'histoire	27
3.1	Le développement historique de la mécanique : avancée formelle ou refonte conceptuelle ?	27
3.2	Les présentations philosophiques et méta-scientifiques de la mécanique	30
4	Les versions de la mécanique	32
2	Formulations de la mécanique classique aujourd'hui : principes et pratique	35
1	Formulation newtonienne	38
1.1	Outils mathématiques de la représentation des systèmes mécaniques	38
1.2	Lois de la dynamique	40
1.2.1	Première loi de Newton : le principe d'inertie	40
1.2.2	Deuxième loi de Newton : le principe fondamental de la dynamique (PFD)	41
1.2.3	Troisième loi de Newton : le principe des actions réciproques ou principe de l'égalité et de la réaction	41
1.3	En pratique	42
1.3.1	Exemples types d'application des lois de Newton	44
1.3.2	Limites pratiques de la formulation newtonienne	49
2	Mécanique analytique – Formulation lagrangienne	50
2.1	Outils mathématiques de la représentation des systèmes mécaniques : espace de configuration et coordonnées généralisées	51
2.2	Principe de d'Alembert et équations de Lagrange	53
2.2.1	Principe du travail virtuel	54
2.2.2	Principe de d'Alembert	55
2.2.3	Les équations de Lagrange. Les concepts d'énergies cinétique et potentielle et de travail	56
2.3	Les principes variationnels. Le principe de Hamilton	58
2.3.1	Déduction du principe de Hamilton à partir du principe de d'Alembert	59
2.3.2	Equivalence du principe de Hamilton et des équations de Lagrange	60
2.4	En pratique	61

2.4.1	Exemples de résolutions de problèmes	62
2.4.2	Limites pratiques de la formulation lagrangienne	66
3	Mécanique analytique – Formulation hamiltonienne	66
3.1	Transformations de Legendre et équations hamiltoniennes du mouvement	67
3.2	Déduction des équations de Hamilton d’un principe variationnel	68
3.3	Transformations canoniques	69
4	En quel sens peut-on dire que les formulations de la mécanique sont équivalentes ?	70
4.1	Équivalence <i>en principe</i>	70
4.2	Bilan des différences	72
4.3	Représentation mathématique et différences pratiques	74
4.4	Les équations de la mécanique comme outils de représentation et d’inférence	78
	Conclusion	82
3	Architecture conceptuelle et versions de la mécanique classique	85
1	L’architecture conceptuelle des formulations de la mécanique classique	88
1.1	La signification des principes dépend de leur place dans l’ar- chitecture déductive de la théorie	88
1.2	Différence architectonique et différence logique	89
1.3	Conception vériconditionnelle et conception inférentielle de la signification	90
2	La notion de version	91
2.1	La notion de version comme hypothèse permettant de rendre compte de l’activité théorique	92
2.1.1	Apprendre la mécanique : développer et enrichir sa version	94
2.1.2	Les versions des experts	95
2.2	Comment étudier les versions de la mécanique ?	98
2.3	Versions et formulations	100
2.4	Versions et histoire de la mécanique	103
3	La version de Lagrange (1788)	106
3.1	L’architecture logique de la <i>Mécanique analytique</i> (1788)	107
3.2	Version de la mécanique et « idéal scientifique » de Lagrange .	113
3.3	Lagrange (1788) et Hamilton (1834) : développement formel contre avancée conceptuelle ?	117
4	Versions de la mécanique lagrangienne	122

4.1	Ernst Mach (1883), <i>La Mécanique. Exposé historique et critique de son développement</i>	125
4.1.1	Ancrer la signification des principes dans une expérience sensible : la défense du concept de force	125
4.1.2	Distinguer entre le contenu physique et la présentation formelle de la mécanique : intelligibilité des principes et économie de la pensée	129
4.2	Heinrich Hertz (1894), <i>Les Principes de la mécanique</i>	134
4.2.1	Images, versions et présentations	134
4.2.2	Les deux images de la mécanique et la version de Hertz : le rejet du concept de force	138
4.2.3	Perfection formelle et intelligibilité de la mécanique	141
4.3	Pierre Duhem (1903b), <i>L'évolution de la mécanique</i>	143
4.4	Une version contemporaine : Cornelius Lanczos (1970), <i>The Variational Principles of Mechanics</i>	147
5	Bilan de l'analyse des versions : différences conceptuelles et différences formelles	151
5.1	Idéal scientifique et « psychologie » des agents	151
5.2	Bilan récapitulatif des versions étudiées	153
5.2.1	Architecture des principes et ordre des concepts	153
5.2.2	Jugements méta-scientifiques et historiques	154
5.3	Changement formel et changement conceptuel	156
	Conclusion	159
4	Que signifie « comprendre la mécanique classique » ?	161
1	Compétences représentationnelles et computationnelles	162
1.1	La compréhension abstraite : Hertz et Duhem	163
1.2	Les deux dimensions de la compréhension située	166
2	Qu'est-ce qui fait qu'une théorie est intelligible ?	169
2.1	Boltzmann et l'évolution des lois de la pensée	170
2.2	Des standards d'intelligibilité aux versions individuelles	173
	Conclusion	174
	Conclusion de la première partie	177
II	Unités et outils d'analyse de l'activité théorique	179
	Vue d'ensemble de la deuxième partie	181

5	Formuler et formaliser : buts et limites des approches formelles des théories scientifiques	185
1	L'empirisme logique et la « signification cognitive » des théories scientifiques	193
1.1	Carnap et la recherche d'un critère de signification cognitive . .	195
1.2	Le dilemme du théoricien	197
1.3	Règles de correspondance et interprétation partielle (Carnap, 1956, 1966)	201
1.4	Le problème des termes théoriques et la solution de Ramsey-Carnap	205
1.5	Ernest Nagel (1961) et l'importance des modèles	209
1.5.1	Les trois composantes de la théorie	210
1.5.2	Les modèles et l'intelligibilité de la théorie : explication des phénomènes et compréhension de la théorie .	214
1.5.3	Le dédoublement de la sémantique	219
1.6	Le statut cognitif des théories	221
1.6.1	Instrumentalisme, réalisme et descriptivisme	221
1.6.2	Le descriptivisme bien compris est un instrumentalisme	222
1.6.3	Les modèles ne disent rien du monde	224
1.6.4	Les théories scientifiques sont toujours déjà interprétées	229
1.6.5	Les formulations de la mécanique comme autant de modèles différents ?	235
1.7	Conclusion : la conception « syntaxique » des théories scientifiques	237
2	La conception sémantique des théories scientifiques	239
2.1	Les théories comme familles de modèles	241
2.2	L'application empirique des théories : modèles logiques et modèles physiques	249
2.3	Le statut cognitif des théories : que nous disent les théories scientifiques ?	253
2.4	À quoi sert la conception sémantique ?	257
3	Formulations et reconstruction formelle : l'identité de la mécanique classique selon les approches formelles	259
3.1	Les vertus de la reconstruction formelle	260
3.2	Formaliser pour aller « au delà » des formulations : conception syntaxique et conception sémantique	264
3.3	La reconstruction formelle suffit-elle à établir l'identité des formulations de la mécanique classique ?	268

Conclusion	274
6 Paradigmes, modèles et gabarits : unités d'analyse globales et locales	279
1 Motivations et limites d'une approche globaliste : la notion de paradigme (Kuhn, 1962/1970)	283
1.1 Le paradigme comme ciment de la communauté scientifique	285
1.1.1 Les composantes principales de la « matrice disciplinaire »	288
1.1.2 Le maintien d'une unité globale	296
1.2 La thèse de l'incommensurabilité des paradigmes et les limites de l'approche globaliste de Kuhn	297
1.2.1 De l'application des généralisations symboliques à l'adoption d'un schème conceptuel	298
1.2.2 Incommensurabilité et réduction inter-théorique	299
1.2.3 Retour à une conception linguistique des théories	303
1.2.4 Le changement scientifique rendu incompréhensible	305
1.3 Paradigmes et versions : deux conceptions de la signification	306
1.3.1 Les paradigmes comme unités monolithiques	307
1.3.2 Sémantique vériconditionnelle et sémantique inférentielle	309
1.3.3 Paradigmes et versions	314
2 Les modèles	315
2.1 Les différents sens du terme « modèle »	317
2.2 Les modèles comme « entités concrètes imaginaires » (Godfrey-Smith, 2006)	319
2.3 L'importance de la forme particulière des représentations	322
3 Unités d'analyse locales : figures et formules	327
3.1 Les cas exemplaires : « Comment les scientifiques relient-ils les expressions symboliques à la nature ? » (Kuhn, 1969)	328
3.1.1 Manipulation et interprétation des généralisations symboliques	329
3.1.2 Invention de nouveaux « formalismes » et invention théorique	333
3.2 Les gabarits : « La syntaxe compte. » (Humphreys, 2004)	337
3.2.1 Les « morceaux concrets de syntaxe »	338
3.2.2 Vers une nouvelle cartographie de l'activité théorique	339
3.2.3 Représentation et calcul : une tension	340

3.2.4	« Réalisme sélectif » (Humphreys, 2004, p. 82)	343
3.3	Figures et formules : des unités conceptuelles	344
3.3.1	Les liens inférentiels entre représentations	344
3.3.2	Différents types de représentation	345
Conclusion	346
7	Outils d'analyse des représentations théoriques	347
1	Systèmes symboliques et types de représentations	349
1.1	La notion de système symbolique	352
1.1.1	Syntaxe et sémantique des systèmes symboliques	352
1.1.2	Quelques propriétés importantes des différents types de systèmes	355
1.2	L'usage épistémique des représentations et leur contenu infor- mationnel	360
1.2.1	Les deux « étapes » de l'utilisation épistémique d'une représentation	361
1.2.2	La détermination contextuelle du contenu des repré- sentations	365
1.3	Vers une typologie des représentations théoriques	367
1.3.1	Représentations linguistiques et représentations spa- tiales	369
1.3.2	Les représentations picturales	372
1.3.3	Les représentations schématiques	375
1.3.4	Les représentations diagrammatiques	381
2	La notion de format de représentation	386
2.1	Quelques exemples de l'importance du format des représentations	389
2.2	Obtenir des informations : coût cognitif et types d'inférences .	396
2.2.1	Contenu informationnel et contenu cognitivement ac- cessible	397
2.2.2	Accessibilité d'une information	399
2.2.3	Vers une définition de la notion de format	404
3	Les formats ne sont pas réductibles à des ensembles de règles	406
3.1	L'expertise comme maîtrise d'un système symbolique ?	407
3.2	Qui maîtrise le système symbolique des diagrammes de Feynman ?	408
3.3	Les équations de la mécanique : le bon algorithme n'existe pas	411
3.4	Renoncer à une définition normative des formats	413
3.5	Versions, formats et activité théorique	414
Conclusion	417

Conclusion de la deuxième partie	418
III Représentations en génétique classique	421
Vue d'ensemble de la troisième partie	423
8 Mendélisme et cytologie : de quoi parle la génétique classique ?	425
1 Les composantes théoriques de la génétique classique	429
1.1 Méthode expérimentale et symbolisme mendéliens	430
1.2 La théorie du gène dans les années 1920	433
1.2.1 La distinction entre génotype et phénotype	434
1.2.2 Les gènes : entités discrètes et stables	436
1.2.3 Les « lois de Mendel »	440
1.2.4 Liaison partielle et crossing-over	443
1.3 La théorie chromosomique de l'hérédité	444
1.3.1 Observations cytologiques et hypothèse chromosomique	445
1.3.2 Janssens (1909) et la chiasmotypie. L'adoption de l'hypothèse chromosomique par Morgan.	447
1.4 Types de représentation en cytologie et en génétique mendélienne	450
2 Identité et versions de la génétique classique selon Philip Kitcher . . .	455
2.1 Le débat sur la réduction et la position de Kitcher (1984) . . .	457
2.1.1 Le contexte du débat sur la réduction de la génétique classique à la génétique moléculaire	457
2.1.2 La pratique de la génétique classique et ses versions selon Kitcher	459
2.1.3 La position anti-réductionniste de Kitcher : l'explication cytologique est préférable.	466
2.2 Critique de la position de Kitcher (1984)	469
2.2.1 Confusion entre schème de raisonnement et niveau d'explication	469
2.2.2 Les « problèmes de pedigree » ne sont pas le cœur du mendélisme	474
Conclusion	476
9 Étude de cas : la cartographie génétique (1913-1939)	479
1 Préhistoire de la cartographie	481
1.1 Le phénomène de liaison partielle	482
1.2 Morgan et les caractères liés au sexe	485

1.3	L'hypothèse du crossing-over	487
2	Sturtevant et la première carte génétique (1913)	491
2.1	La technique d'Alfred Sturtevant	491
2.2	Remarques sur les notions d'observation, de crossing-over et de distance	495
2.2.1	La notion d'observation	495
2.2.2	Le terme de crossing-over	496
2.2.3	La notion de distance	497
3	Le travail du groupe de la drosophile	499
3.1	La carte génétique complète de la drosophile	499
3.2	Vers la confirmation de l'hypothèse chromosomique	501
4	Les années 1930 : confirmation de l'hypothèse chromosomique	504
4.1	Creighton et McClintock (1931) : crossing-over cytologiques et crossing-over génétiques	505
4.2	Painter (1934a,b) : la correspondance des cartes génétiques et des cartes cytologiques	507
5	Les critiques de Goldschmidt (1917), Bateson (1916) et Castle (1919) .	511
5.1	Goldschmidt (1917) : Crossing-over sans chiasmatypie	511
5.2	La critique de Bateson (1916)	513
5.3	Le débat avec Castle (1919) sur la linéarité	514
5.3.1	La critique de Castle (1919)	514
5.3.2	Les réponses de Sturtevant <i>et al.</i> (1919) et Muller (1920) .	523
6	Interprétation de ces débats : que représentent les cartes génétiques ? .	526
6.1	Les cartes génétiques : schémas ou diagrammes ?	527
6.1.1	Les cartes génétiques comme représentations schématiques de chromosomes	528
6.1.2	Les cartes génétiques comme présentation graphique de données numériques	530
6.2	Les cartes génétiques : un nouveau format	532
6.3	Versions du mendélisme : différents « modes d'emploi » des représentations théoriques	535
	Conclusion	537
	Conclusion de la troisième partie	540
	Conclusion générale : les modes d'emploi de la théorie	543
1	Que reste-t-il de la théorie ?	544
2	Différences épistémiques et différences psychologiques : quelle frontière ?	546

3	Compréhension et contenu	550
Annexes		553
A.1	Les mouvements non rectilignes. Moment de force et moment cinétique. Théorème du moment cinétique.	553
A.2	Le mouvement d'un système de points matériels. Forces externes et internes. Théorème du centre de masse.	554
A.3	Les théorèmes de conservation comme conséquences des principes newtoniens	555
A.4	La loi de gravitation universelle	558
A.5	Les différents types de contraintes	558
A.6	Quelques conséquences du principe de d'Alembert	559
A.7	Déduction des équations de Lagrange à partir du principe de d'Alembert	561
A.8	Énergie cinétique, énergie potentielle et travail en mécanique analytique	564
A.9	Déduction du principe de Hamilton à partir du principe de d'Alembert	568
A.10	Déduction des équations de Lagrange à partir du principe de Hamilton.	569
A.11	Théorèmes de conservation et propriétés de symétrie	570
A.12	Le temps comme variable cyclique ; le principe de Jacobi ; le principe de moindre action	572
Bibliographie		576
Index des noms propres		615

Remerciements

J'ai eu la chance, au long des quatre années pendant lesquelles j'ai travaillé à cette thèse, d'être dirigée, soutenue et encouragée avec une attention et une confiance sans égales par mes deux directeurs de thèse, Jean Gayon et Anouk Barberousse. C'est à eux que je souhaite en tout premier lieu exprimer ma gratitude.

Alors que mon projet était encore hésitant, Jean Gayon a accepté de le diriger, et m'a témoigné une confiance dont je lui sais un gré infini. Sa grande exigence et ses conseils m'ont guidée pour donner à mon travail, du moins je l'espère, l'unité qui en fait le sens.

Aucune des étapes de ce travail n'aurait abouti sans Anouk Barberousse. Dès notre rencontre, son enthousiasme, sa disponibilité et sa très grande attention aux questions qui m'intéressaient m'ont aidée à les aborder dans le cadre d'une discipline, dans tous les sens du terme. Quatre ans après, j'admire la manière dont, dès nos premières entrevues, elle a su donner à mon travail l'orientation qui lui a permis d'aboutir. Pour tout cela, je la remercie du fond du cœur.

Je remercie Olivier Darrigol, Paul Humphreys et Pierre Jacob de s'être associés à mes deux directeurs de thèse pour constituer mon jury. Je suis particulièrement sensible au fait que Paul Humphreys ait accepté de faire le voyage depuis la Virginie pour être présent à ma soutenance.

Jacques Dubucs, directeur de l'IHPST, par la place qu'il accorde aux étudiants au sein de ce laboratoire, contribue à en faire un lieu où il fait bon être doctorant. Je tiens à lui en exprimer toute ma reconnaissance.

Outre l'IHPST, mon laboratoire de thèse, et le département d'études cognitives de l'ENS, où j'ai passé la plupart de mes heures de travail, je remercie l'École Doctorale de Philosophie de Paris 1, qui s'est le plus souvent associée à l'IHPST pour financer mes missions. Je remercie également l'UFR de Philosophie de l'Université Paris 1, au sein de laquelle j'ai pu exercer les fonctions de monitrice pendant les trois premières années de ma thèse, et l'UFR de Philosophie de l'Université Paris 4, qui m'a permis de continuer à enseigner pendant ma quatrième année de thèse en tant qu'ATER. Mon séjour de trois mois à la New York University, à l'automne 2007, n'aurait pas

été possible sans, d'une part, le soutien financier de l'Université Paris 1, et, d'autre part, l'hospitalité de NYU, et plus particulièrement de Ned Block, qui m'a invitée à profiter du formidable environnement intellectuel et matériel de son département.

Le soutien et les conseils de Daniel Andler m'ont été très précieux. Je tiens à lui en témoigner ici ma profonde reconnaissance. Je souhaite également remercier Amy Dahan, qui a dirigé mon mémoire de DEA, en 2004-2005, au sein du centre Alexandre Koyré et a su me conseiller et m'aider au moment de choisir un sujet de thèse et le cadre dans lequel j'allais y travailler.

Pour avoir lu et discuté différentes parties de ce travail, ou pour y avoir apporté une contribution indirecte par des échanges qui m'ont permis de formuler certains de mes arguments, je remercie Daniel Andler, Adrien Barton, Ned Block, Julien Boyer, Thomas Boyer, Delphine Chapuis-Schmitz, Amy Dahan, Olivier Darrigol, Patrice David, Steeves Demazeux, Pascal Engel, Michael Esfeld, Roman Frigg, Henri Galinon, Alexandre Guay, John Kulvicki, Soazig Le Bihan, Philippe Mongin, Jim Pryor, Eric Raidl, Julian Reiss et Pierre Wagner. Enfin, Paul Humphreys, Philippe Huneman et Cyrille Imbert, par leurs lectures critiques et les nombreuses discussions que nous avons eues, ont apporté une contribution très substantielle à mon travail. Je les en remercie vivement.

Une pensée, aussi, pour Jean-Michel Muglioni, dont les cours de philosophie, en khâgne, ont été déterminants pour moi. J'espère qu'il saurait reconnaître ici le travail d'une de ses élèves.

Merci aussi à mes amis, docteurs et doctorants, auprès de qui j'ai passé la plupart de mes journées studieuses au DEC, parce que ce travail n'aurait su être le même sans leur compagnie et la possibilité de profiter à chaque instant de leur expertise de logiciens : Henri Galinon en premier lieu, mais aussi Neil Kennedy, Denis Bonnay, Alexandra Arapinis, Julien Boyer et Carlo Proietti. Merci à eux, pour les discussions sérieuses et pour les récréations. À ce même titre, merci à Aude Bandini et Philippe Lusson, qui ont successivement partagé avec moi le bureau du sixième à NYU.

Merci, enfin, à mes proches. À mes amis, et en particulier à Alice Delarbre, qui a littéralement nourri ce travail, et à Delphine de Swardt, qui a su, une fois de plus, trouver les mots justes. À Charlotte Vorms, ma sœur, qui en passant la première les étapes qui conduisent à la soutenance d'une thèse et en m'encourageant à suivre le même chemin, me l'a rendu considérablement moins ardu. À mes parents, Bernard et Eveline Vorms, premières victimes des moments de découragement qui, toujours confiants, m'ont été un soutien moral de chaque instant.

Je les remercie tous très chaleureusement.

Avertissement sur les citations

Sauf mention contraire explicite lors de la première occurrence et dans la bibliographie, la traduction française des ouvrages cités est la mienne. La pagination de référence est celle de l'édition indiquée en bibliographie.

Quand rien n'est précisé, les passages des citations en italique figurent sous cette forme dans le texte original.

Introduction générale : qu'est-ce qu'une théorie scientifique ?

On considère généralement que la connaissance produite par l'entreprise scientifique peut être adéquatement étudiée par la description et l'analyse des théories scientifiques, qui sont censées contenir et exprimer cette connaissance. Un des buts généraux de cette thèse est d'examiner la notion de théorie, et d'évaluer sa pertinence pour l'analyse de la connaissance scientifique. Je me restreins, dans l'ensemble de ce travail, aux théories des sciences empiriques, à l'exclusion, donc, des théories mathématiques.

Dans cette introduction, je présenterai dans un premier temps les présupposés majeurs mis en œuvre dans le travail qui suit, après quoi j'exposerai les principales thèses qui y sont défendues. Un rapide aperçu de l'état des questions qui m'intéressent dans les recherches actuelles me permettra ensuite de situer ma démarche dans le cadre de la philosophie des sciences d'aujourd'hui. Enfin, je présenterai le mouvement d'ensemble de mon argumentation.

1 Présupposés adoptés

1.1 Les théories sont des outils de représentation et d'inférence

On appelle « théorie scientifique » la forme que prend en général la connaissance qui résulte des observations et des expérimentations des scientifiques portant sur un domaine particulier de phénomènes – par exemple, le mouvement des corps macroscopiques. Une théorie se distingue toutefois d'un simple compte rendu d'observation par le fait qu'elle exprime des hypothèses générales à propos des phénomènes qu'elle décrit.

Ces hypothèses consistent à représenter les phénomènes au moyen de concepts

théoriques, dont les référents n'appartiennent pas toujours au domaine de l'observable. Ainsi, la mécanique newtonienne permet de représenter les phénomènes du mouvement au moyen des concepts de masse et de force ; en établissant une relation entre ces deux concepts et les phénomènes observables (selon laquelle la force est le produit de la masse par l'accélération), elle énonce une hypothèse générale à propos de ces phénomènes. C'est cette hypothèse générale qui permet d'aller au delà de la seule description des phénomènes et d'en fournir des explications et des prédictions.

Ces explications et ces prédictions sont le résultat des inférences que les hypothèses théoriques permettent aux agents de tirer. En représentant les phénomènes au moyen des concepts de masse et de force, je peux, en vertu de l'hypothèse selon laquelle la force est le produit de la masse par l'accélération, procéder à un calcul qui me permet de prédire et d'expliquer le comportement des corps que j'étudie. Pour prendre un autre exemple, la théorie mendélienne de l'hérédité, en introduisant le concept de gène et en décrivant le comportement des entités hypothétiques qu'il désigne, permet de prédire et d'expliquer la transmission des caractères observables des individus de génération en génération.

C'est donc en représentant les phénomènes d'une certaine manière – c'est-à-dire, au moyen d'hypothèses mettant en jeu certains concepts – qu'une théorie permet de les prédire et de les expliquer. De cette définition sommaire et consensuelle se dégage le présupposé principal de ma thèse : une théorie est à la fois un outil de *représentation* et un outil d'*inférence* ou de calcul. Ces deux ingrédients sont indispensables à la définition d'une théorie scientifique : si l'une de ces deux fonctions n'est pas remplie, on refuse généralement de parler de théorie scientifique.

D'une part, un discours relatant un ensemble de faits – même vrais – n'est pas un discours théorique s'il n'exprime pas des hypothèses générales permettant de tirer des inférences qui dépassent les seuls faits relatés. À Carlyle, qui, selon Poincaré, aurait écrit « Le fait seul importe ; Jean sans Terre a passé par ici, voilà ce qui est admirable, voilà une réalité pour laquelle je donnerais toutes les théories du monde », le physicien répondrait « Jean sans Terre a passé par ici ; cela m'est bien égal, puisqu'il n'y repassera plus. » (Poincaré, 1902, p. 158) Aussi riche et informatif soit-il, un ensemble d'énoncés exprimant des propositions sur le monde n'est théorique que s'il en offre une représentation telle que l'on puisse, en manipulant cette représentation suivant certaines règles (par exemple, des règles de déduction mathématique), en tirer d'autres propositions.

Inversement, une théorie n'est pas seulement un outil de calcul. Un ordinateur qui, si on y entre certaines données, fournit des résultats justes à propos, par exemple, du temps qu'il fera demain, n'est pas un outil *théorique*, parce qu'il n'est pas un outil de représentation. Le mécanisme qu'il met en œuvre pour accomplir son calcul

n'est pas conceptuel ; ce calcul n'est pas effectué en vertu d'une hypothèse générale à propos des phénomènes. Autrement dit, on n'appellera pas « théorie » un simple instrument de calcul parce qu'on attend d'une théorie qu'elle *nous dise* quelque chose à propos des phénomènes. Une théorie est donc, de façon indissociable, un discours sur le monde et un instrument de calcul.

L'objet d'ensemble de ma thèse est d'explorer certaines conséquences de cette double fonction des théories et de comprendre son articulation. Mon hypothèse de départ est que *c'est en vertu de la manière dont elle représente les phénomènes qu'une théorie permet de tirer des inférences*. Autrement dit, non seulement une théorie est à la fois un discours et un instrument, mais encore c'est en vertu de ce discours et de la manière dont il est présenté qu'elle fonctionne comme instrument. Pour reprendre les termes de Heinrich Hertz, les théories nous fournissent des images des objets du monde extérieur telles que « les conséquences, nécessaires selon la pensée, de ces images soient toujours les images des conséquences, nécessaires selon la nature, des objets reproduits » (Hertz, 1894, p. 67)¹. C'est en disant que les corps ont une masse, qu'ils exercent des forces les uns sur les autres et que leur mouvement est accéléré suivant la proportion évoquée ci-dessus – c'est en fournissant une image « newtonienne » du monde – que la mécanique permet de faire des calculs qui conduisent à l'explication et à la prédiction du mouvement de ces corps.

1.2 Représentation, explication et compréhension

Dans ce qui précède, j'ai situé sur le même plan l'explication et la prédiction des phénomènes observables, considérant qu'elles résultent toutes deux des inférences que nous permettent de tirer les théories, la prédiction concernant les phénomènes à venir, et l'explication les phénomènes qui se sont déjà produits, ou qui se produisent en général. Une précision s'impose sur cet usage de la notion d'explication.

Un des présupposés fondamentaux de la tradition de philosophie des sciences née au début du vingtième siècle des travaux des néopositivistes du Cercle de Vienne est la distinction, formulée par Pierre Duhem (1914), entre l'objet des théories scientifiques et celui de la métaphysique. Si par « explication » on entend désigner l'élucidation des causes premières des phénomènes observables, elles-mêmes inaccessibles aux sens et par conséquent purs objets de spéculation, alors la théorie ne fournit pas, selon Duhem et selon toute la tradition positiviste, d'explications. Elle est « *un système abstrait qui a pour but de RÉSUMER et de CLASSER LOGIQUEMENT un ensemble de*

¹Je cite la traduction française inédite établie par Anouk Barberousse et Alexis Bienvenu. La pagination de référence est celle de l'édition allemande, indiquée en bibliographie.

lois expérimentales, sans prétendre expliquer ces lois. » (Duhem, 1914, p. 3)²

Cela signifie que les concepts que la théorie introduit pour représenter les phénomènes sont des symboles commodes permettant de présenter sous une forme concise et ordonnée l'ensemble des propositions générales tirées des observations empiriques (ce que Duhem appelle les « lois expérimentales »). On peut bien appeler les lois théoriques ainsi formées des « hypothèses », au sens où elles sont les prémisses des inférences permettant de prédire les phénomènes ; elles sont néanmoins nettement distinctes des hypothèses métaphysiques dans la mesure où, selon Duhem, elles ne disent rien de plus que l'ensemble des propositions rapportant des observations empiriques.

Dans cette perspective positiviste, la théorie est bien un outil de représentation et d'inférence, mais les explications qu'elle fournit sont les stricts analogues des prédictions ; elles ont la même forme logique, une forme hypothético-déductive. Quand les positivistes logiques, avec Hempel, réintroduisent la notion d'explication et cherchent à la définir, afin de rendre compte de l'exigence raisonnable selon laquelle une théorie doit fournir une explication des phénomènes de son domaine, c'est sur la base de ce présupposé fondamental. Selon eux, une explication est une forme de description, qui n'exige rien de plus que la connaissance empirique.³

Dans cette thèse, je n'aborderai pas la question de l'explication scientifique ; quand j'emploierai cette notion, ce sera dans le sens pré-analytique et consensuel – dont les positivistes logiques ont cherché à donner une définition formelle –, qui permet de dire qu'une théorie explique les phénomènes qu'elle décrit au moyen d'hypothèses théoriques.⁴ Dire qu'une théorie explique un phénomène signifie donc simplement qu'elle répond à une question de la forme « pourquoi ce phénomène s'est-il produit ? » en permettant de montrer que l'énoncé qui décrit ce phénomène est la conséquence d'autres énoncés décrivant des phénomènes avérés, en vertu d'une hypothèse théorique.⁵ Au-

²Les italiques et petites capitales sont d'origine.

³Pour la conception déductive nomologique de l'explication formulée par Hempel, voir l'article séminal de Hempel et Oppenheim (1948), ainsi que (Hempel, 1965a). Pour les débats sur l'explication scientifique, nés des critiques que la conception de Hempel a suscitées, voir l'ouvrage édité par Salmon et Kitcher (1989), et en particulier la contribution de Salmon (1989).

⁴C'est en ce sens que Carnap affirme qu'il « n'y a pas de véritable opposition entre description et explication. Bien entendu, si l'on entend "description" au sens étroit, comme la pure description de ce que tel scientifique a fait tel jour avec tels instruments, alors les adversaires de la pure description avaient raison [dans les débats qui opposaient, au XIX^e siècle, les partisans de l'explication à ceux de la description] d'exiger davantage, d'exiger une explication réelle. Mais aujourd'hui nous savons bien que la description au sens large, c'est-à-dire l'activité consistant à situer les phénomènes dans le contexte de lois plus générales, fournit le seul type d'explication qui puisse être donné des phénomènes. » (Carnap, 1966, p. 244)

⁵Une théorie explique aussi des phénomènes généraux, en permettant de déduire la loi expérimentale décrivant une régularité empirique d'une hypothèse théorique.

trement dit, une théorie qui permet de déduire la description de certains phénomènes en permet, de ce simple fait, la prédiction et l'explication. On trouvera donc souvent ce terme accolé à celui de prédiction, sans qu'il faille interpréter cela comme une quelconque prise de position dans les débats sur l'explication scientifique.

En revanche, je me démarquerais de la tradition positiviste sur un point important, qui constitue mon second présupposé initial. Outre le rejet de la métaphysique, cette tradition se caractérise par le refus de prendre en considération la dimension psychologique de l'utilisation des théories par les agents. Ainsi, par exemple, une analyse rigoureusement empiriste de l'explication scientifique ne doit pas, dans cette perspective, en appeler au sens psychologique de la notion : le fait qu'un « sentiment de compréhension » soit le résultat d'une bonne explication ne fait pas partie, selon les positivistes logiques, de l'objet du philosophe, mais de celui du psychologue. En conséquence, les positivistes logiques ainsi que leurs héritiers étudient la double fonction des théories (représentative et inférentielle) d'un point de vue logique et non pas d'un point de vue cognitif : ils définissent le contenu des théories comme tout ce que l'on peut en déduire, *en principe*, en vertu de règles objectives, mais ne prêtent pas attention aux processus inférentiels que les agents mettent effectivement en œuvre.

Ma thèse repose sur le parti pris opposé. Je propose d'examiner la double fonction des théories *en adoptant la perspective de leurs utilisateurs*. Par « utilisateurs d'une théorie », j'entends désigner aussi bien les experts qui l'élaborent, la développent et la convoquent pour prédire et expliquer les phénomènes, que les étudiants qui l'apprennent. Adopter cette perspective revient à considérer les théories comme les productions d'esprits humains, et à étudier leur fonctionnement comme outils de représentation et d'inférence sous l'angle de leur utilisation, *en pratique*, par des agents aux capacités cognitives limitées. Autrement dit, je pars du principe qu'il n'y a pas de théorie sans théoricien ou, plus largement, sans des agents qui l'utilisent pour représenter les phénomènes et pour tirer des inférences à leur propos. Cette activité cognitive propre aux utilisateurs d'une théorie, experts ou débutants, est ce que j'appelle l'*activité théorique*. C'est par l'analyse de l'activité théorique que je prétends éclairer la double fonction – représentative et inférentielle – des théories.

Hertz, en affirmant que les théories nous fournissent des images de certains objets du monde en vertu desquelles nous sommes capables de prédire et d'expliquer le comportement de ces objets, prenait acte du fait que, étant « des images de notre esprit particulier », elles « doivent être déterminées aussi par les propriétés de son mode de reproduction des objets » (Hertz, 1894, p. 68). Cela signifie que, puisque c'est en vertu de son fonctionnement comme représentation qu'elle permet de tirer des inférences, une théorie ne remplit pleinement son rôle que pour autant qu'elle est *comprise* par un esprit. Un « système abstrait » de symboles, pour reprendre les

termes de Duhem, n'est véritablement une théorie, dans la perspective que j'adopte, que si le calcul inférentiel opéré sur ces symboles a une dimension *conceptuelle*.

Je pars donc du principe que l'on *comprend les phénomènes* au moyen d'une théorie qui permet de les prédire et de les expliquer si l'on *comprend la théorie elle-même*. Là où les positivistes prétendent définir la fonction prédictive et explicative d'une théorie par l'ensemble de ses conséquences déductives *en principe*, je me propose d'analyser la compréhension des théories, *en pratique*, par les agents.

La compréhension d'une théorie – et donc la compréhension des phénomènes au moyen de cette théorie – repose sur une double compétence : on comprend une théorie quand on est, d'une part, en mesure d'en tirer des propositions prédictives et explicatives à propos des phénomènes, et, d'autre part, quand on comprend *ce qu'elle dit* de ces phénomènes, c'est-à-dire quand on est en mesure de se représenter les phénomènes au moyen de son système de concepts.

Comme je l'ai suggéré précédemment, cette dimension conceptuelle est ce qui distingue une théorie d'une simple machine à calculer : c'est en tant qu'elle est un outil de représentation qu'elle est, aussi, un outil d'inférence. Autrement dit, comprendre comment elle marche – savoir la faire fonctionner comme outil inférentiel – et comprendre ce qu'elle dit – se représenter les phénomènes au moyen de ses concepts – sont deux composantes indissociables de la compréhension d'une théorie, mises en œuvre dans toute activité authentiquement *théorique*.

Cette thèse, dans son ensemble, vise à montrer la fécondité d'une analyse des théories scientifiques qui adopte la perspective des agents. Mon objet central sera donc l'interaction cognitive des agents avec les représentations que leur fournissent les théories. Cela exclut du champ de mon analyse les questions concernant la relation des théories avec les phénomènes, comme celles de la vérité ou de l'adéquation empirique des théories, ainsi que celles relatives aux idéalizations, aux approximations et aux abstractions opérées lors de la description théorique des phénomènes, pour autant, du moins, que ces questions sont abordées sans la prise en compte des agents. Mon entreprise est strictement descriptive, au sens où elle ne vise à énoncer aucune définition normative des théories.

2 Thèses défendues

2.1 La notion de théorie ne permet pas de rendre compte de l'activité scientifique dans la perspective des agents

La première et principale thèse que je défends consiste en une remise en cause de la pertinence de la notion de théorie scientifique, telle qu'elle est habituellement

conçue par les philosophes des sciences. On considère généralement qu'une théorie scientifique, comme, par exemple, la mécanique classique, peut être décrite comme l'expression systématique d'un certain *contenu*, c'est-à-dire d'un ensemble de propositions concernant le monde.

Ce contenu est, la plupart du temps, identifié à l'ensemble des conséquences déductives – ce que les logiciens appellent la *clôture déductive* – d'un petit nombre d'hypothèses, que l'on appelle les « principes » de la théorie. Dans cette perspective, le contenu de la mécanique classique consiste en l'ensemble des propositions que l'on peut en principe déduire des lois fondamentales énoncées par Newton (1687) et reformulées par ses successeurs.

Définir ainsi le contenu d'une théorie revient à faire l'hypothèse selon laquelle la *forme particulière* sous laquelle ses principes sont énoncés n'a pas d'incidence sur son contenu : si les lois de la logique et des mathématiques permettent de montrer l'inter-déductibilité de deux principes, alors ces derniers sont équivalents, puisque leur clôture déductive est identique. L'équivalence logique des principes garantit, en ce sens, l'équivalence de l'ensemble de leurs conséquences empiriques. Une telle conception logico-empirique du contenu des théories caractérise ce que j'appelle, après Paul Humphreys (2004), l'étude de la science *en principe*.

Ce travail vise à montrer que cette perspective est intenable, car elle ne permet pas de rendre compte de la double fonction représentative et inférentielle des théories. Pour en rendre compte, je montre qu'il convient de prêter attention à l'utilisation que les agents font, en pratique, des représentations fournies par les théories. Affirmer cela ne revient pas seulement à proposer, en parallèle d'une étude du contenu des théories, un examen de la science en tant qu'activité humaine : je souhaite montrer qu'une analyse détaillée de l'utilisation des théories rend caduque toute définition strictement logico-empirique de leur contenu.

Pour prendre un exemple que l'on retrouvera dans la suite de ce travail, supposons que l'on dispose d'une équation différentielle permettant, en principe, de prédire le mouvement d'un système physique. Si cette équation n'est pas intégrable, on *ne peut pas* en tirer la fonction décrivant l'évolution temporelle de ce système. Supposons à présent que l'on trouve un moyen de transformer cette équation en une équation soluble, qui permet de tirer les prédictions empiriques souhaitées et, en outre, de déduire plusieurs autres principes débouchant sur tout un ensemble de conséquences insoupçonnées auparavant. Puisque l'on a obtenu la seconde équation à partir de la première, le défenseur d'une conception logico-empirique du contenu des théories affirmera que toutes ces conséquences faisaient déjà partie du contenu de la première équation. L'étude de la science que Paul Humphreys (2004) appelle *en pratique* consiste à refuser une telle affirmation.

Une des premières conséquences de l'adoption de cette perspective est la remise en cause de toute distinction ferme entre la forme et le contenu des théories. En effet, la modification formelle d'un principe peut avoir des conséquences importantes sur ce qu'il est possible, en pratique, d'en déduire. Or, il est fréquent – le cas de la mécanique classique en est un exemple – qu'une « même » théorie puisse être exprimée sous des formulations différentes. En conséquence, la remise en cause de la distinction de principe entre forme et contenu implique, à son tour, de décomposer la notion traditionnelle de théorie et d'étudier dans le détail les *processus inférentiels* mis en œuvre par les agents qui utilisent des hypothèses théoriques *sous une forme particulière*.

Pour montrer la fécondité de cette perspective, mon argumentation reposera en très grande partie sur l'analyse d'exemples ; en outre, par l'examen critique de certaines tentatives de définition logico-empirique du contenu des théories, je montrerai qu'elles manquent des aspects essentiels de l'activité théorique. Cela me conduira donc à renoncer à une conception des théories scientifiques conçues comme des « tous » abstraits, c'est-à-dire indépendants de l'utilisation que les agents peuvent en faire, et à définir d'autres notions plus à même, selon moi, de rendre compte de l'activité théorique.

2.2 Une hypothèse théorique : la notion de version

La première de ces notions est celle de *version*. Elle correspond, pour ainsi dire, au versant *cognitif* de ce que l'on appelle traditionnellement « théorie ». Prenant acte de la nécessité de rendre compte des processus inférentiels mis en œuvre par les agents, je ferai l'hypothèse d'un ensemble de représentations mentales propre à chaque individu utilisant une théorie – en entendant par « théorie » un ensemble d'hypothèses théoriques à propos d'un domaine de phénomènes.

Cela signifie que je supposerai que chaque agent, au cours de son apprentissage et de sa pratique scientifiques, se forme un ensemble de représentations mentales au moyen desquelles il raisonne à propos des phénomènes. Cet ensemble de représentations mentales, naturellement, est extrêmement contraint par les relations logiques entre les concepts de la théorie. Cependant, chaque individu, en fonction de son histoire, de l'apprentissage qu'il a suivi, de la forme sous laquelle les hypothèses théoriques lui ont été présentées et de l'ordre dans lequel elles ont été introduites, relie différemment ces représentations entre elles et au monde empirique. Quoique leur architecture déductive soit contrainte par des relations objectives entre les concepts, on peut supposer que de nombreux *chemins inférentiels* différents restent possibles.

Dans mon argumentation, l'introduction de la notion de version a le statut d'une

hypothèse théorique permettant de rendre compte de l'activité scientifique du point de vue des agents. Cependant, dans la mesure où leurs représentations mentales ne nous sont pas accessibles – elles sont des entités hypothétiques destinées à rendre compte de leur vie cognitive et, plus précisément, de leurs raisonnements –, elles ne peuvent pas être étudiées « directement ».

Par conséquent, mon analyse portera essentiellement sur les écrits et les propos *explicites* de certains scientifiques, ainsi que sur les représentations *externes* que les agents construisent et manipulent au cours de leur apprentissage et de leur pratique scientifiques. Introduire la notion de version revient à considérer ces représentations externes comme la source et le produit d'un ensemble de représentations mentales. Afin d'étudier la manière dont les agents les utilisent dans leurs raisonnements, je propose des outils d'analyse, parmi lesquels la notion de format de représentation.

2.3 Un outil d'analyse : la notion de format

Renonçant à fonder l'étude de l'activité théorique sur la notion trop abstraite de théorie, et opérant un changement de perspective qui consiste à étudier les produits de l'activité théorique du point de vue des agents, je montrerai qu'il convient, pour étudier le détail de la représentation et de l'inférence scientifiques, de prêter attention aux représentations *particulières* utilisées par les agents, et non aux théories prises comme des « tous ». Cette démarche implique de s'intéresser à *tous les types* de représentations (externes) utilisées dans la pratique scientifique. Je serai ainsi conduite à prendre en compte non seulement les représentations linguistiques, qu'il s'agisse d'énoncés en langage naturel ou d'équations mathématiques, mais encore les autres types de support utilisés pour exprimer, développer et appliquer des hypothèses théoriques : images picturales, diagrammes, schémas, graphes, etc.

Afin de rendre compte de la diversité de ces types de représentation, et de son incidence sur les raisonnements des agents, je définirai et j'analyserai la notion de *format de représentation*. Cette notion vise à rendre compte du fait que deux représentations dont le contenu est en principe identique (par exemple, une liste de données numériques et la présentation de ces mêmes données sous la forme d'un tableau ou d'un graphe) ne facilitent pas les mêmes inférences chez les agents. Autrement dit, leur contenu, quoique équivalent en principe, n'est pas également *accessible* aux agents. Or, dans la perspective que je défends, c'est l'étude du contenu cognitivement accessible des représentations qui est pertinent pour une analyse de l'activité théorique. La recherche d'une définition précise de la notion de format – et, corrélativement, de celle du contenu cognitivement accessible d'une représentation – me permettra de montrer la chose suivante : dès lors que l'on prête attention aux conséquences de

la *forme* d'une représentation *pour les raisonnements des agents*, on est conduit à prendre en compte, non seulement cette forme, mais encore les capacités cognitives des utilisateurs de cette représentation, ainsi que le but épistémique dans lequel elle est utilisée.

Dans cette analyse, comme dans celle de la notion de version, je m'appuie sur une conception *inférentialiste* du contenu des représentations. On peut, à gros traits, distinguer deux manières concurrentes de définir le contenu d'un concept (ou la signification d'un terme). La première consiste à le définir en termes de conditions de vérité : le contenu d'un concept ou plutôt d'un énoncé dans lequel il figure est alors déterminé par l'état du monde qui le rendrait vrai. La seconde affirme que le contenu d'un concept est, au moins partiellement, déterminé par le rôle ou la fonction qu'il remplit dans les inférences des agents. Cette dernière – que l'on appelle « conception inférentialiste » – me semble la plus adaptée pour rendre compte du contenu des représentations théoriques. Sans entrer dans les débats difficiles opposant ces différentes conceptions de la signification, j'adopte donc et mets en œuvre cette perspective ; j'espère ainsi en montrer la fécondité pour une analyse de la connaissance scientifique.

3 État de la question et méthodologie

Ce travail, consistant en une critique de la conception de la connaissance scientifique issue des travaux des positivistes logiques du Cercle de Vienne et de leurs héritiers, se situe, de ce simple fait, dans la lignée de ces travaux. C'est en effet en prenant au sérieux les motivations et les buts de ces philosophes que je prétends en faire la critique ; un chapitre (le chapitre 5) est ainsi consacré à un examen des approches formelles des théories, et vise à montrer que ma démarche trouve sa source dans le constat de tensions *internes* à ces approches.

Je chercherai à montrer que ces tensions sont dues à l'adoption, par les positivistes logiques et leurs héritiers, du présupposé que j'entends rejeter : celui selon lequel on peut – et on doit – étudier la connaissance scientifique en faisant abstraction des conditions dans lesquelles elle est construite, établie et utilisée.

En outre, les outils que je mets en œuvre et les analyses qui m'ont directement influencée proviennent tous de travaux qui se situent dans l'héritage « large » de l'empirisme logique et de ses critiques. Je me propose de les présenter rapidement dans ce qui suit.

3.1 L'héritage kuhnien

L'œuvre de Thomas Kuhn est sans doute celle à la lecture de laquelle mon travail doit le plus. *La Structure des révolutions scientifiques*, parue en 1962, se présente comme une critique radicale de la philosophie des sciences des positivistes logiques. Elle repose sur le rejet de la plupart des présupposés fondamentaux de ces derniers, et en particulier de l'entreprise qui consiste à reconstruire la connaissance scientifique sans prêter attention au contexte de son élaboration, aux pratiques qui lui sont associées, et aux agents qui la construisent.

Les thèses de Kuhn, aussi riches que contradictoires, ont fait l'objet d'une quantité innombrable de critiques. Je n'entends pas, dans ce travail, en proposer la défense ou en prolonger la critique. Comme je le montrerai dans un chapitre qui lui est en partie consacré (chapitre 6), mon travail se démarque du sien sur des points importants. S'il en est un héritier, c'est moins pour ses thèses célèbres sur l'incommensurabilité des paradigmes que pour la partie de son travail consacrée à l'étude des pratiques de ce qu'il appelle la « science normale ».

J'adopte en effet, et je mets en œuvre, un des présupposés fondamentaux de la pensée de Kuhn, celui selon lequel l'analyse de la connaissance scientifique passe par une étude de la manière dont elle est *apprise* par les agents. Mon analyse de l'activité théorique, comme celle de Kuhn, vise en effet à en donner une définition qui inclue aussi bien les processus de l'apprentissage que ceux du développement et de l'application des théories. À son insistance sur les *communautés scientifiques*, je substituerai toutefois une attention aux *agents cognitifs individuels*.

Je rejette en revanche, comme cela apparaîtra clairement dans le choix de mes études de cas, un autre des présupposés kuhnien : celui selon lequel il existe une frontière nette entre le développement de la science normale, c'est-à-dire celle des pratiques normées des sciences arrivées à maturité et institutionnellement reconnues (la science des « disciplines » scientifiques), et les changements que Kuhn appelle « révolutionnaires », qui séparent les différentes traditions de science normale. L'essentiel de mon propos est en effet de remettre en cause l'aspect « monolithique » des théories les mieux reconnues, et de montrer qu'il existe d'authentiques différences conceptuelles *au sein* d'une « même » théorie.

3.2 Les philosophes de la science « en pratique »

Plus récemment, plusieurs philosophes des sciences, que l'on peut décrire, pour certains, comme les héritiers du tournant pragmatique engagé par les travaux de Kuhn, ont proposé des analyses de la connaissance scientifique reposant sur le rejet de plusieurs présupposés des positivistes logiques. Ainsi, Nancy Cartwright (1983) et

ses élèves ont souligné l'inadéquation de toute conception hypothético-déductive des théories scientifiques, selon laquelle l'explication et la prédiction des phénomènes empiriques consistent en une *application* déductive des théories. Prenant comme objet premier d'étude les processus de *modélisation*, c'est-à-dire la construction, la modification et la manipulation de modèles, conçus comme des outils partiellement indépendant des théories, ils conçoivent la science comme une activité plus inventive et imaginative que les reconstructions formelles des théories ne le suggèrent.

D'autres philosophes, proches à certains égards de cette perspective, comme par exemple Ian Hacking, et surtout Philip Kitcher, se situent plus clairement dans l'héritage kuhnien, et décrivent, plutôt que le contenu logico-empirique des *théories*, les « styles » (voir par exemple Hacking, 1982, 1983a,b, 1992a,b, 2003) ou encore les « schèmes » (Kitcher, 1981, 1983, 1993) de raisonnement propres à différentes disciplines et traditions scientifiques.

Si mon travail est proche, par plusieurs aspects, de ceux de ces auteurs, il vise cependant à prêter une attention plus précise aux processus cognitifs effectivement mis en œuvre par les agents au cours de leur apprentissage et de leur pratique scientifiques. Comme cela apparaîtra dans la critique que je proposerai de la notion de modèle au chapitre 6, ainsi que dans la comparaison de mon étude de la génétique classique avec celle que propose Kitcher (1984) (chapitre 8), les analyses de ces auteurs se concentrent, selon moi, sur des objets (modèles, schèmes de raisonnement) encore trop abstraits pour permettre une analyse détaillée de la représentation et de l'inférence scientifiques.

Parmi les philosophes des sciences que l'on peut classer dans cette catégorie grossièrement définie comme celle de la « science en pratique », Paul Humphreys (2004) est sans aucun doute celui auquel mon travail doit le plus. La critique qu'il propose des approches formelles des théories dans son ouvrage *Extending Ourselves* repose, entre autres, sur le constat de l'importance de la forme particulière – qu'il appelle la « syntaxe » – des équations et des représentations manipulées dans l'application empirique des théories. Une partie du chapitre 6 de ce travail est consacrée à l'examen d'un des outils d'analyse – la notion de *gabarit* – qu'il propose dans ce but.

La différence majeure entre la perspective de Humphreys et la mienne est la dimension cognitive que je donne à l'étude de l'activité théorique. Alors que son propos est de rendre compte de l'importance croissante de l'utilisation des ordinateurs dans ce qu'il appelle la « science computationnelle », mon but est de rendre compte de la compréhension individuelle que les agents, du débutant à l'expert, ont des représentations théoriques qu'ils utilisent. En conséquence, dans mon travail, la frontière entre la science *en principe* et la science *en pratique* ne passe pas précisément au même endroit que chez Humphreys.

3.3 Les approches cognitives de l'activité scientifique

Une seconde branche de l'héritage kuhnien – après la science en pratique – est celle de l'étude cognitive de l'activité scientifique. Plusieurs philosophes et historiens des sciences (par exemple Nersessian, 1984, 2008; Thagard, 1992), inspirés par les propos de Kuhn concernant l'apprentissage scientifique et le changement conceptuel, cherchent dans les sciences cognitives et en particulier la psychologie expérimentale (entre autres, Carey, 1985; Spelke, 1991; Carey, 1991; Gopnik, 1996) des hypothèses permettant d'analyser le raisonnement scientifique.

L'examen de l'interaction cognitive des agents avec les représentations qu'ils utilisent gagnerait sans doute à la mobilisation de certains résultats des recherches en sciences cognitives, comme ceux auxquels font référence ces auteurs. Cela permettrait en particulier d'approfondir l'analyse des notions de version et de format, par une compréhension de la relation entre les représentations mentales des agents et les représentations externes. Quoique fortement inspiré par la lecture de certains travaux portant sur ces questions, mon travail reste cependant, autant que possible, agnostique quant aux hypothèses cognitives permettant d'étayer mon propos. Étant donné le caractère encore neuf et rarement consensuel de la plupart des hypothèses existantes sur le format des représentations mentales et les processus computationnels qui implémentent les raisonnements des agents, je reste aussi prudente que possible dans mon utilisation de ces hypothèses.

3.4 Les études sur la science (*science studies*)

Critiquer le positivisme logique en insistant sur la pratique de la science, comme je le fais dans cette thèse, peut sembler rapprocher mon propos d'une autre tradition d'analyse de l'activité scientifique – qui se réclame d'ailleurs, elle aussi, de l'héritage kuhnien –, celle des « études sur la science » (*science studies*) ou « études sociales de la science » (*social studies of science*). Ce champ, fort divers, est né dans les années 1970 d'un rejet radical de tous les présupposés du positivisme et, plus largement, du refus de la plupart des catégories classiques d'analyse de la connaissance scientifique, comme les notions de rationalité, d'objectivité, de vérité et de méthode (voir, entre autres Bloor, 1976; Latour, 1995; Callon et Latour, 1991; Pickering, 1992).

Considérant que la science est une activité humaine dont le contenu est déterminé par les interactions sociales de ses acteurs, les différents défenseurs de cette approche prétendent rejeter tout discours normatif sur les sciences, et étudier la science « telle qu'elle se fait ». Cela les conduit souvent à adopter une méthodologie fondée sur des *études de cas* historiques (voir par exemple Shapin et Schaffer, 1985) ou sociologiques, et à prêter une grande attention aux pratiques et aux « traces » concrètes produites

par les scientifiques et circulant entre eux (voir par exemple Latour et Woolgar, 1979).

Dans la mesure où ma démarche repose, elle aussi, sur une attention à la pratique et aux représentations concrètes utilisées par les scientifiques, ainsi que sur une méthodologie favorisant les études de cas – comme cela apparaît clairement à la vue du plan général de ma thèse, qui accorde deux parties sur trois à de telles études – il convient de marquer nettement ce qui, au delà de ces points communs, différencie ma perspective de celle des études sur la science.

Tout d’abord, cette thèse étant un travail de philosophie et non d’histoire ou de sociologie des sciences, les études de cas que j’y développe s’appuient, en très grande partie sur des sources secondaires ayant déjà « fait le travail » de l’enquête historique. En particulier, la première partie de ma thèse, consacrée à la mécanique classique, n’est pas une enquête sur l’histoire de la mécanique, et je ne prétends rien ajouter au savoir historique concernant cette théorie. Ma seconde étude de cas, consacrée à la cartographie génétique dans les années 1920, peut sembler davantage influencée par les méthodes des études sur la science. Cependant, l’intérêt que je porte aux pratiques des agents et à l’utilisation de représentations concrètes est motivé par des questions d’une nature différente de celles qui guident les études sur la science.

En effet, c’est à la pratique *en tant qu’elle fait partie de l’activité théorique* des agents que je m’intéresse, et non aux pratiques en tant qu’elles sont insérées dans un réseau d’interactions sociales, institutionnelles, et économiques. Étudier l’interaction cognitive des agents avec les représentations théoriques qu’ils utilisent est, pour moi, le moyen d’analyser le *contenu conceptuel* des connaissances produites, et non ce que les spécialistes des études sur la science appellent la « construction des faits scientifiques ».

En conséquence, tout en m’appuyant parfois sur des études de cas issues de travaux appartenant à cette tradition (comme, par exemple, celle de David Kaiser, 2005 sur les diagrammes de Feynman), je rejette certains présupposés intellectuels et méthodologiques qui la caractérisent souvent. En particulier, le présupposé qui consiste à refuser toute pertinence aux catégories d’objectivité, de vérité et de rationalité me semble aller à l’encontre d’une analyse féconde de l’activité théorique qui, indépendamment de la question de savoir si elle débouche effectivement sur l’établissement de croyances vraies, est, pour une très grande part, guidée par la recherche d’une telle connaissance. Le parti pris qui découle de ce présupposé, et qui consiste à n’analyser *que* les pratiques, sans prêter aux sujets de ces pratiques l’intention d’apprendre et de comprendre, me semble limiter la fécondité de ces analyses.

En conséquence, j’adopte le présupposé inverse : m’intéressant à la compréhension des agents, je pars du principe qu’« il y a bien quelque chose à comprendre ». Admettre que cette compréhension est toujours *située*, et enquêter sur les conséquences de ce

constat, ne revient pas à faire du contenu des connaissances scientifiques un pur produit des pratiques et des interactions sociales.⁶

4 Plan de l'argumentation

Cette thèse s'ouvre et se ferme par une étude de cas. La première, qui occupe les quatre premiers chapitres, est consacrée à la mécanique classique, et me permet de soulever les différentes questions que je prétends traiter. La seconde, qui en occupe la troisième et dernière partie, est destinée, par l'examen d'un exemple tiré des sciences biologiques (et plus précisément de la génétique classique) à mettre en œuvre la méthode proposée et les outils d'analyse élaborés pour répondre à ces questions.

4.1 Première partie – Formulations et versions de la mécanique classique

La première partie de ce travail, des chapitres 1 à 4, est consacrée à un exemple paradigmatique de théorie scientifique, celui de la mécanique classique. Cette partie dans son ensemble vise deux objectifs. Le premier est de montrer, par l'exemple, la première des thèses que je défends, annoncée précédemment : la notion de théorie, conçue comme un ensemble d'hypothèses jointes à leur clôture déductive, ne permet pas de rendre compte de l'utilisation des théories comme outils de représentation et d'inférence. Le deuxième est d'introduire la notion de version, et d'en montrer la fécondité pour l'analyse de plusieurs problèmes soulevés par l'examen de la mécanique classique.

Après un **premier chapitre** introductif présentant ces différents problèmes, j'aborde l'étude de la mécanique sous deux perspectives différentes, synchronique et diachronique.

Dans le **chapitre 2**, je m'intéresse à la mécanique classique telle qu'elle est enseignée et utilisée *aujourd'hui*. Cette perspective anhistorique me permet de montrer que l'identité de ce que l'on considère habituellement comme une seule et même théorie est fortement remise en cause par une attention à son utilisation pour représenter et résoudre des problèmes particuliers. En effet, les différentes formulations sous lesquelles la mécanique peut être exprimée et utilisée sont autant d'outils de représentation et d'inférence différents. Cette analyse me permet de définir et de

⁶La fermeté de ma prise de distance vis-à-vis des études sur la science ne doit pas masquer, d'une part, la grande diversité des thèses qui y sont défendues – auxquelles j'ai consacré un mémoire de DEA – ni, d'autre part, la richesse de certaines études de cas historiques, comme, par exemple, celle de David Kaiser (2005) sur laquelle une partie du chapitre 7 du présent travail s'est largement appuyée.

justifier le choix d'une étude du contenu des connaissances théoriques centrée sur la pratique.

Le **chapitre 3**, en introduisant la notion de version, poursuit l'examen de la mécanique classique dans la perspective des agents. Cet examen me permet d'éclairer deux types de problèmes. En premier lieu, je montre qu'il existe, entre les différentes formulations de la mécanique telle qu'elle est enseignée et utilisée aujourd'hui, d'authentiques différences conceptuelles, dont une analyse logico-empirique des théories ne permet pas de rendre compte. Ce constat, qui motive l'introduction de la notion de version, me permet, en second lieu, d'éclairer un problème d'ordre historiographique : aborder l'étude de certains épisodes du développement historique de la mécanique sous l'angle de la notion de version me permet en effet de remettre en question la possibilité d'une distinction tranchée entre le développement formel d'une théorie et sa refonte conceptuelle.

Enfin, le **chapitre 4** propose un bilan des analyses qui précèdent à la lumière de la notion de compréhension.

4.2 Deuxième partie – Unités et outils d'analyse de l'activité théorique

La deuxième partie de ce travail (chapitres 5 à 7) est consacrée d'une part à l'examen critique des analyses disponibles de la connaissance scientifique, et d'autre part à la précision et la justification de ma propre démarche, ainsi qu'à l'élaboration de mes outils d'analyse.

Dans le **chapitre 5**, je présente et je critique deux grandes tentatives pour étudier le contenu des théories scientifiques en faisant abstraction des conditions dans lesquelles elles sont construites, formulées, utilisées et comprises : celle mise en œuvre par les empiristes logiques, héritiers des travaux du Cercle de Vienne, puis celle des défenseurs de la conception dite « sémantique » des théories, qui se présente comme une critique de la première. Le but de ce chapitre est double. En premier lieu, l'analyse de ces approches qui reposent sur le présupposé inverse du mien – une étude de la science en principe contre une étude de la science en pratique – me permet, par contraste, de préciser mon approche. En second lieu, je procède à une critique interne de ces entreprises, en montrant que leur refus de prendre en compte les utilisateurs des théories sont la cause de plusieurs tensions au sein même de leur propos. Ce faisant, je justifie l'adoption du parti pris méthodologique consistant à se placer dans la perspective des agents.

Le **chapitre 6** est consacré à l'examen de certaines analyses qui se présentent elles-mêmes comme des critiques des approches formelles, et dont la mienne est l'hé-

ritière. Ayant renoncé à centrer l'étude de l'activité théorique sur les théories elles-mêmes, je cherche dans ces analyses des outils permettant de rendre compte de l'interaction cognitive des agents avec les représentations qu'ils utilisent dans l'apprentissage et la pratique scientifiques. Une telle étude doit satisfaire plusieurs exigences. Elle doit en particulier permettre de rendre compte de l'importance de la forme de ces représentations, et du fait qu'elles acquièrent leur signification en tant qu'elles sont utilisées par les agents comme outils de représentation et d'inférence. J'examine successivement plusieurs notions destinées à remplacer ou à compléter celle de théorie, parmi lesquelles celle de paradigme, proposée par Kuhn (1962/1970). Par une analyse critique de ces notions, je précise ma propre perspective. Je suis conduite à défendre la thèse selon laquelle il convient d'étudier tous les types de représentation – et non pas seulement les représentations linguistiques – et de renoncer à en étudier le contenu sur le modèle de la référence linguistique.

Le **chapitre 7** élabore les outils d'analyse permettant de rendre compte de la diversité des représentations utilisées dans la pratique scientifique. J'y propose, en m'appuyant sur les travaux de Nelson Goodman (1968/1976), l'ébauche d'une typologie de ces représentations, qui sera mobilisée aux chapitres 8 et 9. J'y introduis ensuite la notion de format, destinée à rendre compte de l'importance de la forme des représentations pour les raisonnements des agents. Initialement motivée, comme je l'ai mentionné précédemment, par le constat du fait que deux représentations ayant le même contenu peuvent faciliter différents types d'inférences, l'analyse de la notion de format me permet de montrer que la prise en compte des capacités cognitives des agents dans l'analyse du contenu des représentations implique de reconsidérer la définition de ce contenu. Cela me permet de compléter la remise en cause, amorcée dès la première partie, de la distinction entre la forme et le contenu des représentations.

4.3 Troisième partie – Représentations en génétique classique

La dernière partie de ce travail (chapitres 8 et 9) est consacrée à une seconde étude de cas, en génétique classique. Elle vise deux objectifs principaux. Le premier est de montrer que l'approche de l'activité théorique défendue dans les deux parties précédentes n'est pas uniquement applicable aux sciences physiques. En effet, l'examen des pratiques représentationnelles des scientifiques permet selon moi de saisir des aspects essentiels de toute activité théorique. Il permet, en outre, d'apporter un éclairage nouveau sur certaines questions conceptuelles propres à un domaine scientifique particulier. Le second objectif principal de cette partie est ainsi d'examiner, au moyen des outils élaborés dans les deux premières parties, certains problèmes relatifs à la nature et au statut des différentes hypothèses qui constituent ce que l'on appelle

« génétique classique ».

Le **chapitre 8** vise à isoler les différentes composantes théoriques de la génétique classique en les identifiant à l'utilisation de certains types de représentation. J'établis une distinction, au sein de la génétique classique, entre ce que j'appelle le mendélisme « pur » et l'étude des processus cellulaires (cytologie). Cela me permet de montrer les limites de l'approche de Philip Kitcher (1984), elle-même motivée par le constat de l'inadéquation des conceptions classiques des théories, et de suggérer que certaines questions concernant les relations inter-théoriques sont susceptibles d'être éclairées par une analyse des représentations utilisées dans la pratique scientifique.

Le **chapitre 9** met cette suggestion en œuvre dans une étude détaillée de la construction, de l'utilisation et de l'interprétation des cartes génétiques dans les années 1920. Les cartes génétiques incarnent en effet l'articulation des deux domaines théoriques distingués au chapitre 8. Leur construction a fait l'objet de plusieurs débats au sein de la communauté des généticiens dans les années 1920. Je montre que l'interprétation que les différents généticiens font des cartes génétiques est partiellement guidée par leurs engagements théoriques et manifeste ce que je propose de considérer comme leurs versions du mendélisme. Cela me permet de montrer que les représentations théoriques ne sont pas tant les moyens d'accéder à une théorie dont elles seraient l'expression, que le lieu même de l'activité théorique.

Première partie

Formulations et versions de la mécanique classique

Vue d'ensemble de la première partie

Cette partie, consacrée à la mécanique classique, vise deux objectifs principaux. Le premier est de soulever plusieurs questions relatives à la notion de théorie, et de montrer, par l'analyse détaillée d'un exemple canonique de théorie scientifique, la première des thèses que je défends dans ce travail : si l'on prête attention aux détails de l'utilisation d'une théorie comme outil de représentation et d'inférence, on s'aperçoit que l'identité de ce que l'on considère habituellement comme une seule et même théorie est remise en question. Les différentes formulations de la mécanique classique, telle qu'elle est enseignée et utilisée aujourd'hui, n'offrent pas la même représentation des phénomènes, et ne facilitent pas les mêmes inférences.

Le deuxième objectif principal de cette partie est d'introduire la notion de version et d'en montrer la fécondité par l'analyse de plusieurs réflexions épistémologiques et méta-scientifiques que la mécanique a inspirées, au cours de son développement, chez certains des principaux acteurs de son histoire. En présentant ces réflexions comme l'expression d'autant de versions de la mécanique, je montre que la notion de version permet d'enrichir celle de théorie et de rendre mieux compte, par la place qu'elle accorde aux agents cognitifs, de la double dimension, représentationnelle et inférentielle, de l'activité théorique.

Cette partie a un double statut : étude de cas mettant en œuvre la démarche que je défends, elle sert aussi d'introduction aux différents problèmes que l'ensemble de ce travail vise à résoudre. Qu'on me pardonne, en conséquence, l'apparente disparité des questions que j'y soulève ; la suite de ce travail, en y proposant des réponses, contribuera à en dégager l'unité. Pour l'heure, je procède à une présentation de la mécanique classique sous plusieurs aspects qui, pris ensemble, fournissent une image (partielle) de la manière dont cette théorie est enseignée, utilisée et comprise aujourd'hui. Cette présentation, outre l'occasion de soulever les questions qui m'intéressent, me fournit une réserve d'exemples dans laquelle je puiserai tout au long de ce travail.

Un premier chapitre introductif offre une cartographie des principaux problèmes

abordés dans la partie et propose une définition initiale des notions qui y seront analysées. Dans le chapitre 2, je présente les principales formulations de la mécanique classique telles qu'elles sont enseignées et utilisées aujourd'hui, en montrant que leur équivalence *en principe* est remise en question par l'analyse de leur utilisation, *en pratique*, comme outils de représentation et d'inférence. J'insiste particulièrement sur le lien, au sein de chacune de ces formulations, entre la manière dont elles représentent les phénomènes et les inférences qu'elles permettent de tirer. Le chapitre 3 est consacré à l'analyse de quelques versions de la mécanique – à ne pas confondre avec ses formulations – qui ont été exprimées au cours de l'histoire. Il vise à remettre en question l'idée d'une distinction nette entre différence conceptuelle et différence formelle. Enfin, le chapitre 4 dresse un premier bilan des problèmes identifiés, à la lumière de la notion de compréhension.

Chapitre 1

Quel est le contenu de la mécanique classique ?

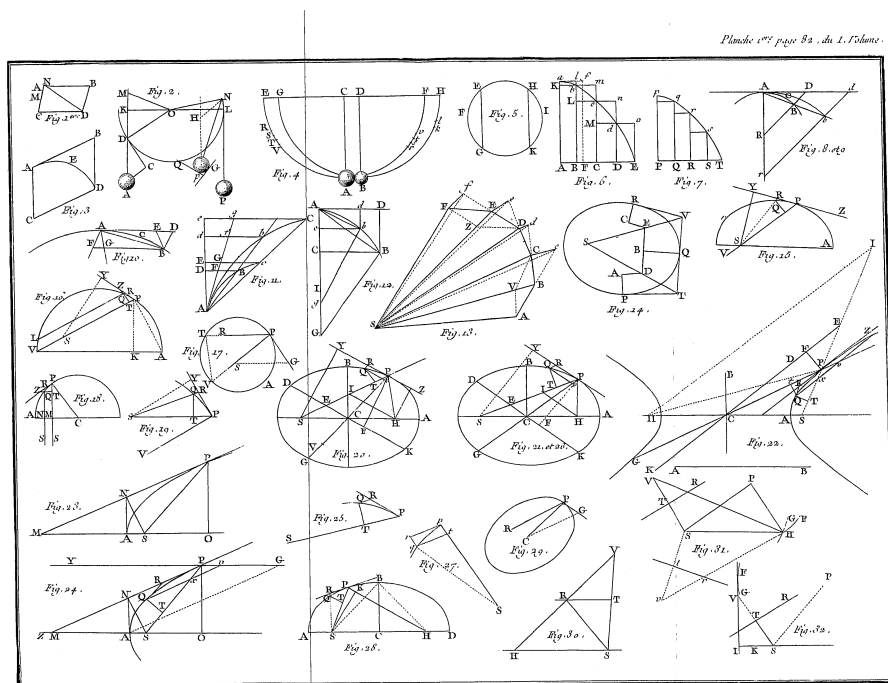


Planche de figures extraite des *Principia* de Newton (1687), « Du mouvement des corps », livre premier.

Quand on cherche à définir ce que l'on entend par « théorie scientifique », il semble naturel de se référer à des exemples de théories existantes. La mécanique classique en est un des plus fameux, et a longtemps été considérée comme un canon de théorie scientifique. Dans ce chapitre introductif, je présente sommairement la manière dont on identifie habituellement cette théorie, par la définition de son domaine d'application et l'énoncé des principes et théorèmes au moyen desquels elle permet de prédire et d'expliquer les phénomènes de ce domaine. J'expose ensuite les principales difficultés qu'une telle définition rencontre. Ces difficultés feront l'objet des analyses détaillées des chapitres suivants, qui justifieront et mettront en œuvre l'approche de l'activité théorique que j'entends défendre.

1 Identifier le contenu de la mécanique classique

Une première manière d'identifier la mécanique classique consiste à définir le domaine de phénomènes auquel elle s'applique et dans les limites duquel elle est valide : c'est la théorie physique qui étudie le mouvement des corps macroscopiques, dans les cas où leur vitesse est très inférieure à celle de la lumière, ce qui exclut les phénomènes quantiques et relativistes.¹ Des phénomènes comme la chute des corps sur terre, le mouvement oscillatoire d'un pendule, la trajectoire relative des astres, sont typiquement dans le domaine d'étude de la mécanique.

À l'époque de leur première présentation systématique par Isaac Newton (1687) dans les *Philosophiae Naturalis Principia Mathematica*, les principes de la mécanique étaient conçus comme fournissant une explication universelle des phénomènes du mouvement. De plus, pour Newton, ainsi que, plus tard, pour les défenseurs de l'hypothèse atomiste, la mécanique était censée expliquer non seulement les phénomènes du mouvement, mais aussi l'ensemble des phénomènes physiques et chimiques. Au vingtième siècle, la double révolution constituée par l'énoncé des théories de la relativité par Albert Einstein et la naissance de la mécanique quantique assigne des limites au domaine d'application de la mécanique désormais « classique » ; on sait maintenant que sa validité n'est pas universelle, puisqu'elle est fautive en ce qui concerne les phénomènes microscopiques et cosmologiques.

Définir son domaine d'application comme je viens de le faire n'est pas suffisant pour identifier une théorie ; on peut en effet avoir plusieurs théories différentes, contradictoires ou incompatibles entre elles, portant sur un même domaine de phénomènes. On veut donc aussi pouvoir identifier *ce que la théorie dit* de ces phénomènes, c'est-à-

¹Cette définition du domaine de la mécanique classique est quelque peu restrictive, dans la mesure où l'on peut, au moyen des équations de la mécanique, étudier des comportements plus généraux que le mouvement des corps. C'est cependant à eux seuls que je m'intéresserai dans ce travail.

dire l'ensemble des hypothèses au moyen desquelles elle nous permet de les prédire et de les expliquer. Cette exigence semble facile à satisfaire pour la mécanique classique : elle peut être présentée, comme c'est le cas dans les *Principia* de Newton, sous la forme d'un ensemble déductif d'énoncés exprimant des hypothèses théoriques. Leur présentation systématique attribue à ces différentes hypothèses un statut plus ou moins fondamental. Quoique les usages varient, les plus générales et fondamentales sont couramment appelées « principes² ». Viennent ensuite les lois et les théorèmes. Les lois correspondent à la forme que prennent les principes quand ils sont appliqués à certains types particuliers de mouvements (par exemple, la loi de gravitation ou la loi du pendule). Les théorèmes peuvent être définis comme des conséquences remarquables des principes (par exemple, les théorèmes de conservation). Une telle présentation systématique de la mécanique est si commune que c'est souvent sur une référence à cet exemple que s'appuient l'analyse et la définition philosophiques des notions de principes, lois et théorèmes.

Ces différentes hypothèses, jointes à l'ensemble de leurs conséquences déductives, semblent constituer un ensemble clairement défini que l'on peut appeler le *contenu* de la mécanique. Pourtant, un examen attentif des différentes manières dont ce contenu peut être exprimé fait surgir plusieurs difficultés.

2 Première difficulté : les différentes formulations contemporaines de la mécanique classique

Aujourd'hui, la mécanique classique est enseignée et utilisée sous plusieurs *formulations* différentes. On distingue généralement la formulation dite « newtonienne » ou encore « vectorielle³ » de la formulation dite « analytique » ou « variationnelle », laquelle se présente elle-même sous plusieurs formulations différentes (lagrangienne, hamiltonienne, théorie de Hamilton-Jacobi). Si le nom de chacune de ces formulations indique que la paternité en est attribuée à des savants d'époques différentes qui ont contribué aux développements de la mécanique (Newton, 1687; Lagrange, 1788; Hamilton, 1834), le premier ensemble de problèmes que je souhaite soulever gagne à être abordé en faisant abstraction du développement historique qui a conduit à instituer ces formulations comme standard.

Ces problèmes sont ceux de la coexistence, *aujourd'hui*, dans l'enseignement et la pratique de la mécanique, de plusieurs formulations, qui diffèrent non seulement par

²Newton les appelle « axiomes ». On les appelle aussi parfois les « lois de Newton ».

³La formulation vectorielle de la mécanique n'est en fait pas due à Newton, mais elle utilise les mêmes concepts physiques (en particulier celui de force). Je reviendrai sur ce problème plus loin.

les systèmes de coordonnées et la forme des équations utilisées pour représenter et expliquer les phénomènes du mouvement, mais aussi par leurs principes et concepts fondamentaux. Insistons sur un point de vocabulaire : dans toute la partie consacrée à la mécanique classique (chapitres 1 à 4), je réserverai le terme de *formulation* aux formes standard sous lesquelles la mécanique est aujourd'hui enseignée et utilisée, en ignorant délibérément leur genèse au cours de l'histoire. Il conviendra de bien distinguer ce terme de ceux de « présentation » et de « version », que j'introduis plus loin dans le présent chapitre.

Les formulations contemporaines de la mécanique, malgré les différences que je viens de mentionner – systèmes de coordonnées, forme des équations, principes et concepts –, sont habituellement considérées comme équivalentes : une chaîne de déductions mathématiques permet de passer de l'une à l'autre, ce qui semble garantir que leurs conséquences empiriques sont strictement identiques. Ainsi, les prédictions fournies par les différentes formulations de la mécanique n'entrent jamais en contradiction entre elles, même pour un domaine restreint de phénomènes. La définition logique des concepts qui y figurent est la même, même si leur place dans l'architecture d'ensemble de la théorie change : le concept de force, fondamental dans la formulation newtonienne, devient secondaire dans les formulations variationnelles, qui s'appuient sur le concept d'énergie, mais sa définition n'est pas modifiée.

La relation entre les formulations de la mécanique classique est donc très différente, par exemple, de la relation entre la mécanique newtonienne et la théorie einsteinienne de la relativité, qui ne donnent pas la même définition du concept de masse, et dont les prédictions diffèrent pour les très grandes vitesses. Pour toutes ces raisons, on admet généralement que le contenu de la mécanique n'est pas affecté par la diversité de ses formulations, considérées comme différentes expressions d'*une seule et même théorie*.

Je prétends cependant montrer que la notion de théorie qui sous-tend une telle affirmation n'est pas satisfaisante, car elle ne rend pas compte de la manière dont la mécanique fonctionne comme outil de représentation et d'inférence. Si l'on prête attention à la manière dont on utilise les équations issues des différentes formulations de la mécanique pour résoudre des problèmes concrets, on s'aperçoit que leur équivalence logique ne garantit pas l'identité de la théorie *en tant qu'elle est utilisée et comprise par des agents*. En effet, les différentes formulations de la mécanique n'offrent pas la même représentation des phénomènes et ne facilitent pas les mêmes inférences. Là où la mécanique newtonienne parle de forces, et représente ces dernières par des vecteurs, la mécanique hamiltonienne parle d'énergie et utilise des quantités scalaires. En outre, comme on le verra, selon le cas que l'on étudie, certaines formulations sont, en pratique, inutilisables, quand d'autres permettent d'atteindre aisément la solution

du problème.

Le chapitre 2, en insistant sur ces différences entre les formulations contemporaines de la mécanique, vise à montrer qu'il convient de renoncer à identifier le contenu d'une théorie à l'ensemble de ses conséquences déductives, abstraction faite des agents qui les utilisent. Pour cela, je m'efforcerai de mettre en valeur la distinction formulée par Paul Humphreys (2004), entre une étude de la science *en principe*, au regard de laquelle l'équivalence des formulations de la mécanique, et par conséquent l'identité de son contenu, si elle ne sont pas triviales, semblent à peu près assurées, et une étude de la science *en pratique*, qui consiste à accorder une place importante à ce que l'utilisateur d'une théorie peut en faire, et qui invite à réexaminer ces notions d'équivalence et de contenu.

3 Deuxième difficulté : les présentations successives de la mécanique au cours de l'histoire

Afin de bien distinguer les différents problèmes auxquels je m'intéresse, je propose d'appeler « présentation » de la mécanique tout exposé de ses principes, qu'il soit destiné à en remanier en profondeur l'appareil mathématique, à en examiner ou en clarifier l'organisation conceptuelle, ou encore simplement à exposer ces principes afin de les enseigner et de les utiliser. Une présentation est donc un objet concret et public historiquement situé, qu'il s'agisse des *Principia* de Newton (1687), de la *Mécanique analytique* de Lagrange (1788), de *La Mécanique* de Mach (1883), des *Principes de la mécanique* de Hertz (1894), d'un manuel quelconque ou encore de l'exposé du chapitre 2 du présent travail. Par contraste, les formulations contemporaines de la mécanique peuvent être décrites comme des classes d'équivalence de présentations : on peut proposer plusieurs présentations différentes de chacune de ces formulations, et ces dernières n'entrent pas dans un rapport de succession historique ni ne portent la marque de l'époque à laquelle le physicien à qui en est attribuée la paternité les a pour la première fois présentées.

3.1 Le développement historique de la mécanique : avancée formelle ou refonte conceptuelle ?

Il existe plusieurs types différents de présentations de la mécanique. D'abord, les présentations des « grands noms » de l'histoire de la mécanique, comme celles de Newton (1687), Lagrange (1788) et Hamilton (1834), font partie d'un vaste ensemble

de travaux⁴ qui, du dix-septième au dix-neuvième siècles, ont contribué à construire et à développer l'appareil mathématique de la mécanique et, par des généralisations successives, à en étendre le domaine d'application et dresser des ponts entre elle et d'autres branches de la physique. Après leur première formulation systématique par Isaac Newton (1687), les principes de la mécanique ont ainsi fait l'objet de plusieurs reformulations et réorganisations, dont certaines ont été jugées suffisamment innovantes par la postérité pour que le nom de leur auteur devienne le label d'une formulation standard.

Il convient cependant de distinguer soigneusement entre les présentations historiquement données par Newton (1687), Lagrange (1788) et Hamilton (1834) et les présentations que l'on donne aujourd'hui des formulations dites « newtonienne », « lagrangienne » et « hamiltonienne ». Par exemple, la formulation newtonienne d'aujourd'hui est exprimée en langage vectoriel et au moyen d'équations différentielles, alors que toutes les démonstrations des *Principia* sont faites en langage géométrique et non à l'aide du calcul différentiel. De même, la formulation lagrangienne d'aujourd'hui est fondée sur le principe de Hamilton, alors que l'auteur de ce principe est âgé de 10 ans au moment de la mort de Lagrange.

Ce constat conduit à s'interroger sur les raisons au nom desquelles on peut affirmer que les formulations newtonienne et lagrangienne méritent ou non de porter ces noms, c'est-à-dire les raisons qui motivent le jugement selon lequel Lagrange, par exemple, est ou non le père de la formulation lagrangienne. Expliciter ces raisons implique de mettre au jour ce que l'on considère constituer le contenu de ces différentes formulations ; on pourra alors discriminer entre ce qui leur est essentiel et ce qui relève simplement des aspects contingents de leur présentation. On pourrait, par exemple, affirmer que l'utilisation du calcul différentiel, quoique commode, n'est pas essentielle à l'expression du contenu de la formulation newtonienne : on jugerait alors que la réécriture différentielle des *Principia* (opérée par Varignon, 1700a,b)⁵ est une simple avancée formelle, et non une refonte conceptuelle de la mécanique de Newton.

Distinguer ainsi entre une avancée formelle et une authentique refonte conceptuelle est une opération particulièrement délicate dans le cas de la mécanique. Habituellement, une manière simple d'opérer une telle distinction est de repérer des différences logiques et empiriques : on juge, par exemple, que la théorie de la relativité générale est conceptuellement différente de la théorie newtonienne de la gravitation parce que, en donnant (entre autres) une définition différente du concept de masse, elle conduit à des prédictions empiriques qui contredisent les prédictions

⁴Ceux, entre autres, de Varignon (1700a,b), d'Euler (1744), de d'Alembert (1743), de Laplace (1843-1846), de Poisson (1838) et de Jacobi (1837, 1842).

⁵Pour la « réécriture » des *Principia* par Varignon, voir (Blay, 1992, 1993, 1995, 2002).

newtoniennes.

Les cas typiques de changement scientifique étudiés par les philosophes et historiens des sciences sont ainsi des changements *de théorie* : les deux théories n'ont pas le même contenu, et cela se manifeste sans ambiguïté dans les définitions qu'elles donnent de leurs concepts fondamentaux et dans les prédictions empiriques auxquelles elles conduisent. En ce sens, la frontière entre un développement formel et un changement conceptuel semble nette, le changement conceptuel correspondant à un changement logico-empirique. Des conceptions très différentes du changement scientifique s'accordent d'ailleurs sur ce point : que l'on conçoive les relations inter-théoriques sur le modèle de la réduction, comme les empiristes logiques (voir en particulier Nagel, 1961, chap. 11)⁶, ou sur le modèle de l'incommensurabilité, comme Kuhn (1962/1970)⁷, les différences conceptuelles sont la plupart du temps identifiées à des différences logico-empiriques. L'analyse des empiristes logiques permet certes aussi de rendre compte de changements intra-théoriques, mais les conçoit uniquement comme les conséquences de la confrontation à de nouvelles données empiriques, dont il s'agit de rendre compte en modifiant les hypothèses de la théorie en place – parfois de façon suffisamment mineure pour que l'on ne considère pas que cela donne naissance à une nouvelle théorie. Il s'agit donc, ici encore, d'un changement logico-empirique.

Le développement historique de la mécanique classique ne peut pas être décrit ainsi. En effet, les présentations successives des « grands noms » de la mécanique, si elles sont motivées par la volonté d'en permettre l'application à des problèmes nouveaux et d'en approfondir les relations avec d'autres théories, n'impliquent pas de changement logico-empirique ; leur moteur n'en est pas la découverte d'une nouveauté empirique inexplicable dans le cadre théorique en place et dont il s'agirait de rendre compte en modifiant certaines hypothèses. Le développement de la mécanique, intimement lié aux avancées mathématiques dont il est à la fois le moteur et la conséquence, consiste en une réécriture mathématique et en une réorganisation de ses principes. S'il débouche, dans certains cas, sur la définition de concepts nouveaux comme celui d'action, cela ne modifie pas celle des concepts existants.

⁶On dit qu'une théorie T_1 est réduite par une théorie T_2 quand tous les phénomènes expliqués et prédits par T_1 sont prédictibles et explicables par T_2 , c'est-à-dire quand T_1 est une conséquence logique de T_2 . En outre T_2 peut corriger certaines prédictions de T_1 , faisant alors apparaître les prédictions de T_1 comme des approximations des prédictions de T_2 .

⁷Pour Kuhn, deux théories portant sur le même domaine de phénomènes sont dites incommensurables dès lors que certains de leurs concepts centraux, même s'ils sont exprimés par les mêmes termes, reçoivent des définitions logiquement incompatibles. Je présenterai et critiquerai les thèses de Kuhn au chapitre 6.

Si l'on applique à ce cas les modèles habituels du changement scientifique, l'histoire de la mécanique est celle d'un développement purement formel. En ce sens, aucune nouveauté conceptuelle n'est introduite par Lagrange ou Hamilton ; en ce même sens, la mécanique présentée par Hamilton est équivalente à celle de Newton. Ce jugement sur l'histoire des sciences est la contrepartie diachronique de l'affirmation de l'équivalence (synchronique) entre les formulations actuelles héritées des présentations successives de la mécanique. Il revient à nier la possibilité d'un changement *intra-théorique*⁸ qui soit authentiquement *conceptuel*.

J'affirme que la notion de théorie qui sous-tend un tel jugement est, ici encore, insatisfaisante. En identifiant le contenu d'une théorie à l'ensemble des conséquences déductives de ses principes, et en dressant une frontière nette entre développement formel et innovation conceptuelle, on manque d'authentiques différences théoriques. Montrer cela sera un des objets du chapitre 3 ; l'introduction de la notion de version, en donnant une place centrale à la compréhension que les utilisateurs d'une théorie en ont, permettra une analyse plus fine de la question du changement théorique que celle fondée sur la notion classique de théorie, ainsi que la remise en question d'une distinction nette entre développement formel et avancée conceptuelle.

3.2 Les présentations philosophiques et méta-scientifiques de la mécanique

Un deuxième type de présentations de la mécanique retiendra mon attention. Outre les profonds remaniements opérés par les « grands noms » de l'histoire de la mécanique, cette théorie a fait l'objet, en particulier à la charnière des dix-neuvième et vingtième siècles, de plusieurs examens critiques destinés à en clarifier l'ordre des principes et la signification des concepts.

Dans un contexte où les prémices de la crise de la physique qui va donner naissance aux théories de la relativité et à la théorie quantique⁹ se font sentir, marquant la fin du règne de la physique classique, physiciens et mathématiciens¹⁰ s'interrogent et débattent sur le contenu et la place dans la connaissance scientifique de la mécanique classique, longtemps considérée comme *la* théorie physique toute entière, dont les autres disciplines devaient être des branches. Les explications fournies par la mécanique ont en effet longtemps été érigées en standard d'intelligibilité : les physiciens

⁸Pour une analyse de la notion de changement intra-théorique sur la base de ce même exemple, voir l'article d'Anouk Barberousse (2008), auquel l'analyse du chapitre 3 doit beaucoup.

⁹Henri Poincaré (1902, 1905) est un des premiers à diagnostiquer cette crise comme telle.

¹⁰Entre autres, Hertz (1894), Mach (1883), Poincaré (1902, 1905), Boltzmann (1897), von Helmholtz (1847) et Kirchhoff (1877).

considéraient qu'une description partiellement incompatible avec les lois de la mécanique était inintelligible.¹¹ Autour de 1900, l'universalité de ce modèle explicatif est remise en cause. Il devient urgent, pour le défendre ou pour le combattre, d'en avoir une conception parfaitement claire. Cela passe, la plupart du temps, par la recherche d'une présentation de la mécanique qui permettrait de faire apparaître clairement le statut de ses différentes hypothèses et de donner une définition précise à ses différents concepts.¹²

La distinction entre ces présentations, que j'appelle « philosophiques », ou « méta-scientifiques », et celles des « grands noms » de l'histoire de la mécanique, est affaire de degré, et non de nature. Certes, les auteurs dont il est question ici écrivent à la fin du dix-neuvième siècle, à une époque où les développements de la mécanique classique ont contribué à lui donner la forme sous laquelle on la présente aujourd'hui ; son développement est, pour ainsi dire, achevé. Cependant, non seulement ce type de réflexions méta-scientifiques se trouve également dans les écrits des principaux acteurs de l'histoire de la mécanique, mais encore la recherche d'une présentation des principes de la mécanique qui en fasse apparaître le plus clairement possible le contenu peut donner lieu à des réécritures qui gagnent parfois le statut de nouvelles formulations. Par exemple, la présentation proposée par Hertz (1894) se donne à la fois comme une clarification conceptuelle et la proposition d'une nouvelle formulation ; la plupart des manuels approfondis de mécanique d'aujourd'hui y consacrent un développement (voir par exemple Lanczos, 1970, pp. 130-132).

Ces réflexions philosophiques et méta-scientifiques présentent plusieurs intérêts pour mon propos. D'abord parce qu'elles sont autant d'exemples de présentations de la mécanique, dont l'analyse détaillée conduit à remettre en cause la possibilité d'isoler et de s'accorder sur un ensemble d'hypothèses qui formerait le socle de l'édifice de la mécanique.

Ensuite parce, selon la manière dont chacun de ces auteurs procède en cherchant à identifier et à présenter clairement le contenu de la mécanique, il exprime du même coup une certaine conception des théories et de leur contenu. Par exemple, tout en reconnaissant que les théories fournissent des images mentales qui portent la marque des esprits qui les produisent, Hertz (1894) semble orienter sa recherche vers un idéal qui consisterait à distinguer clairement le contenu d'une théorie de sa forme, jugée contingente.

¹¹Pour une analyse de la manière dont les physiciens ont cherché à sauver les modèles explicatifs de la physique classique dans le cadre de la révolution quantique, voir (Chevalley, 1991).

¹²Ainsi Poincaré regrette-t-il que « les traités de mécanique ne distinguent pas bien nettement ce qui est expérience, ce qui est raisonnement mathématique, ce qui est convention, ce qui est hypothèse. » (Poincaré, 1905, chap. VI, « La mécanique classique », p. 111)

Enfin, la manière dont chacun de ces auteurs traite le problème de la diversité des formulations de la mécanique et de son développement historique me fournit autant de matériau pour analyser ces problèmes. Autrement dit, ces auteurs, par leurs réflexions, me fournissent un objet d'analyse, en ajoutant à la liste des présentations de la mécanique qui en ont été historiquement proposées, et sont en même temps des interlocuteurs qui se trouvent confrontés au problème qui est le mien. La méthode que je mettrai en œuvre pour les lire consistera à traiter ces présentations comme les expressions d'autant de *versions* de la mécanique, dont je propose à présent une définition sommaire, que le chapitre 3 contribuera à étayer en montrant la fécondité pour l'étude de l'activité théorique.

4 Les versions de la mécanique

Chacune des présentations de la mécanique, qu'il s'agisse de la *Mécanique analytique* de Lagrange (1788), de *La Mécanique* de Mach (1883), des *Principes de la mécanique* de Hertz (1894), ou encore d'un manuel de référence contemporain comme ceux de Goldstein (1950/2002) ou de Lanczos (1970), propose une certaine perspective sur la mécanique, par la manière dont elle en ordonne les différents principes et décrit leur articulation, dont elle définit les concepts qui y figurent, dont elle évalue la différence entre les formulations et, le cas échéant, la contribution des précédents acteurs de l'histoire de la mécanique. Ces différentes perspectives correspondent à autant de manières de *comprendre* la mécanique ; après Anouk Barberousse (2008), je propose d'appeler la manière de comprendre la mécanique propre à chaque individu, sa « *version* » de la théorie. Je la définis initialement comme l'ensemble des représentations mentales qu'il se forme au cours de son apprentissage et de sa pratique de cette théorie.

La notion de version me permet d'émettre une hypothèse contribuant à donner une image unifiée de l'activité théorique, cette dernière expression servant à désigner non seulement les processus d'invention et de développement d'une théorie, mais aussi celui de son apprentissage. Je suppose en effet que chacun des utilisateurs de la mécanique, du débutant à l'expert, en a une certaine version, plus ou moins riche, plus ou moins robuste ; elle dépend à la fois de l'apprentissage qu'il a suivi, de ses connaissances d'arrière-plan, ainsi que de ses talents et des types de raisonnement auxquels il est habitué. Elle détermine la manière dont il utilise les différents principes et concepts de la mécanique pour représenter les phénomènes du mouvement et pour tirer des inférences lorsqu'il résout des problèmes.

Mon propos, il convient de le préciser, n'est pas d'affirmer que la version d'un débutant est une version de la mécanique classique au même titre que celle d'un expert

ou, encore moins, que celle de Hamilton. En premier lieu, il existe bien une différence entre les inférences autorisées par les principes et les équations de la mécanique, et les inférences fautives ; les versions des agents, pour autant qu'elles comptent comme des versions de la mécanique, sont donc extrêmement contraintes par les relations logiques entre ses concepts.

En second lieu, à supposer qu'un débutant fasse un usage correct des principes qui lui sont enseignés, sa version se distingue de celle d'un expert, entre autres, par son inefficacité à résoudre la plupart des problèmes que la mécanique permet de résoudre. En outre, à la différence de celles de Lagrange ou de Hamilton, par exemple, la plupart des versions ne sont pas innovantes : elles sont le résultat de l'apprentissage et de l'utilisation d'un ensemble d'hypothèses qui ont été formulées par d'autres, mais ne débouchent pas elles-mêmes sur une manière inédite de formuler ces hypothèses.

Enfin, à la différence de celle de Lagrange ou de Hamilton, mais aussi de Hertz, de Mach, et de tous les auteurs de manuels, la version d'un profane ne fait généralement pas l'objet d'une présentation *explicite*. C'est aux versions de la mécanique qui ont été exprimées, au moins partiellement, dans des présentations explicites, que je vais m'intéresser au chapitre 3. Étudier les différentes présentations de la mécanique qui en ont été données au cours de son histoire comme l'expression des versions de leurs auteurs m'offrira un moyen d'aborder les différents problèmes que j'ai identifiés dans ce chapitre et de mettre en évidence les limites d'une conception des théories qui ne prend pas en compte leur utilisation par les agents.

L'examen des différentes *formulations actuelles* de la mécanique (chapitre 2), puis celui des *versions* qui en ont été exprimées dans certaines *présentations* (chapitre 3) contribuera à remettre en question l'idée d'un contenu objectif clairement isolable de la mécanique, indépendant des différentes manières dont il est formulé et compris. On verra ainsi que les définitions et les critères habituellement invoqués pour identifier une théorie scientifique semblent se dissoudre à l'examen de l'exemple même qui est censé en être un des canons.

Chapitre 2

Formulations de la mécanique classique aujourd'hui : principes et pratique



Le professeur Tryphon Tournesol et son pendule.
Dessin de Hergé (1948), *Les Sept Boules de cristal*,
Casterman.

Ce chapitre est consacré à la présentation de trois formulations standard de la mécanique classique, la formulation newtonienne d'une part, et les formulations analytiques (lagrangienne et hamiltonienne) d'autre part, telles qu'elles sont enseignées et utilisées aujourd'hui. J'ignore délibérément, dans cette présentation, le développement historique qui a conduit à les instituer comme formulations standard, et je n'interroge pas le bien fondé de leurs appellations. Ce qui m'intéresse au premier chef, dans ce chapitre, est la coexistence et l'utilisation, à la même époque, de plusieurs formulations différentes d'une théorie. Je me suis principalement appuyée, pour cet exposé, sur deux ouvrages de référence en mécanique, ceux de Herbert Goldstein (1950/2002) et de Cornelius Lanczos (1970). Les cours de Jean-Michel Raimond (2000), ainsi que les articles de Jeremy Butterfield (2004) et de Jill North (2009a,b) m'ont également été utiles.

Le but de ce chapitre est de montrer que, si l'on conçoit une théorie comme un outil de représentation et d'inférence, et si l'on analyse en détail la manière dont elle remplit ces deux fonctions, les critères habituels d'identité des théories sont remis en cause ; en particulier, je montrerai que la définition du contenu d'une théorie comme l'ensemble de ses conséquences déductives – la clôture déductive de ses principes – devient intenable. Cela me permettra de justifier l'adoption, pour l'analyse de l'activité théorique, d'une perspective centrée sur les agents.

Les différentes formulations de la mécanique sont habituellement considérées comme équivalentes, en vertu de l'inter-déductibilité de leurs principes. Ce jugement repose sur une notion d'équivalence que j'appellerai « *logique* », selon laquelle deux ensembles d'énoncés expriment le même contenu si l'ensemble de leurs conséquences déductives est le même. On va voir que, à supposer même que l'on puisse démontrer que les formulations de la mécanique satisfont des critères d'équivalence logique très stricts¹, il reste d'importantes différences entre elles, qu'une analyse de l'activité théorique doit prendre en compte.

Ces différences concernent la manière dont les formulations de la mécanique fonctionnent comme outils de représentation et comme outils d'inférence. Du point de vue

¹L'équivalence logique de deux ensembles d'énoncés dépend, entre autres, du langage dans lequel on définit cette équivalence. Pour une analyse classique de l'équivalence des différentes formulations d'une théorie dans la tradition empiriste logique, voir par exemple l'article de Quine (1975). Quine vise à donner un critère d'identité des théories (par delà leurs différentes formulations) qui soit plus faible que celui de la stricte équivalence logique, mais plus fort que celui de la seule équivalence empirique. La stricte équivalence logique ne permettrait pas, en effet, de s'accommoder des cas de formulations qui n'emploieraient pas les mêmes termes mais qu'un simple échange de termes suffirait à rendre équivalentes (imaginons, par exemple, les formulations d'une même théorie en anglais et en français). Cependant, le critère recherché par Quine doit permettre de garantir l'équivalence empirique des formulations *a priori*, quel que soit l'état du monde.

de la représentation, elles diffèrent par leurs principes et leurs concepts fondamentaux, ainsi que par les systèmes de coordonnées et la forme des équations utilisées pour représenter la dynamique des systèmes physiques. C'est sur ce second ensemble de différences – les différences de représentation mathématique – que le présent chapitre est centré. Le fil directeur de mon exposé est la relation entre les outils mathématiques utilisés pour *représenter* les phénomènes du mouvement et leur utilisation, en pratique, pour *résoudre des problèmes* de mécanique, c'est-à-dire pour prédire et expliquer l'évolution dynamique de certains systèmes physiques. Comme on le verra, les différences de représentation mathématique ont d'importantes conséquences sur les inférences que l'on peut tirer au moyen des équations. Mon but est d'examiner la manière dont la représentation des phénomènes au moyen de ces langages mathématiques joue sur les raisonnements et la compréhension qu'en ont les utilisateurs de la théorie.² Je ne ferai que mentionner, dans l'exposé des trois formulations, les différences concernant les concepts et les principes, sur lesquelles je reviendrai au prochain chapitre.

Résoudre un problème de mécanique consiste, typiquement, à trouver les *fonctions* qui décrivent l'évolution de la position et de la vitesse du système étudié au cours du temps. Ces fonctions sont les solutions d'un ensemble d'*équations différentielles*³ qui gouvernent la dynamique du système ; cela signifie qu'elles disent quelles trajectoires le système peut suivre en le représentant dans un espace donné et à l'aide d'un certain système de coordonnées. La résolution d'un problème de mécanique suit typiquement deux étapes. La première est celle de la *mise en équation* : il s'agit d'écrire

²Je contournerai donc la difficile et importante question de savoir comment et pourquoi les structures des théories mathématiques permettent de décrire et de prédire efficacement le comportement des phénomènes empiriques. Cette question, en étroite relation avec la question du réalisme scientifique – un des problèmes étant de savoir si, comme l'affirment les défenseurs du réalisme structural, les mathématiques permettent de représenter les structures réelles, physiques du monde – fait aujourd'hui l'objet de débats profonds et complexes (voir, par exemple, Shapiro, 1983; Resnik, 1997; Ladyman, 1998), dont je ne m'occuperai pas ici.

³Le rôle fondamental des équations différentielles en physique a été maintes fois souligné ; ainsi, Arnold (1988, préface) affirme : « La découverte fondamentale de Newton, celle qu'il a considéré devoir garder secrète et qu'il a publiée seulement sous la forme d'une anagramme, est la suivante : *Data aequatione quocunque fluentes quantitae involvente fluxiones invenire et vice versa*. En langage mathématique contemporain, cela signifie : "Il est utile de résoudre des équations différentielles". » Jeremy Butterfield (2004, p. 19) cite Poincaré : « Sans la physique, nous ne connaîtrions pas les équations différentielles ». Voir aussi (Mach, 1883, p. 321) : « La physique s'habitue progressivement à considérer la description des faits au moyen d'équations différentielles comme son véritable but ». Pour une analyse historique et philosophique des notions de fonctions et du rôle des équations différentielles dans la solution des problèmes dynamiques, voir (Butterfield, 2004, p. 10-23) et les sources historiques auxquelles il renvoie (Bottazini, 1986; Lützen, 2003; Kline, 1972; Youschkevitch, 1976).

les équations du mouvement du système étudié à partir des informations dont on dispose à propos de ce système. La deuxième étape est la *résolution de ces équations*, c'est-à-dire leur intégration. On va voir que les différences de représentation mathématique entre les trois formulations de la mécanique ont des conséquences pratiques très importantes sur les deux étapes de la résolution de problèmes mécaniques.

Après avoir présenté successivement les trois formulations de la mécanique (sections 1, 2 et 3) et mis en évidence que chacune d'elles offre des outils de représentation et d'inférence différents, je dresserai le bilan de ces différences et j'approfondirai, à la section 4, l'analyse de la relation entre un changement dans la forme mathématique des équations et une modification des raisonnements que les agents doivent faire pour utiliser ces équations dans un calcul. Cela me conduira à réexaminer les notions d'équivalence et de contenu au nom desquelles on affirme généralement que les formulations de la mécanique sont les expressions d'une seule et même théorie, et de montrer qu'elles sont insatisfaisantes pour une analyse de l'activité théorique qui cherche à rendre compte de la double fonction représentative et inférentielle qu'une théorie doit remplir.

1 Formulation newtonienne

La formulation newtonienne⁴ est la première, et souvent la seule, que l'on apprend avant d'entreprendre des études approfondies de physique ou d'ingénierie. Elle est souvent introduite comme un préalable aux autres formulations, dont l'appareil mathématique est moins élémentaire. Dans ce qui suit, je présente les outils mathématiques qu'elle utilise pour représenter les phénomènes mécaniques (section 1.1), puis les principes qui permettent d'obtenir les équations du mouvement gouvernant la dynamique des systèmes (section 1.2), et enfin la manière dont, en pratique, on peut utiliser ces équations pour résoudre des problèmes (section 1.3).

1.1 Outils mathématiques de la représentation des systèmes mécaniques

Un des concepts de base de la mécanique newtonienne est celui de *point matériel* (ou *particule*). Il s'agit d'un point géométrique, c'est-à-dire dépourvu d'étendue mais

⁴On l'appelle aussi « vectorielle ». Précisons cependant que c'est Euler qui en 1749 écrit la deuxième loi de Newton sous la forme qu'on lui connaît aujourd'hui (en coordonnées cartésiennes, cependant, et non à l'aide d'une notation vectorielle). En outre, la notion de point matériel, et l'interprétation atomiste qui lui est associée, est attribuée à Boscovitch, dans sa *Philosophiae naturalis Theoria* de 1758.

doté d'une grandeur appelée « masse » représentée par un réel positif m .⁵ L'*espace de représentation du point matériel* en mécanique newtonienne est un espace affine euclidien, défini par un espace vectoriel de dimension finie (et plus précisément à trois dimensions, quand le point matériel se meut dans un espace physique à trois dimensions) muni d'un produit scalaire, permettant de définir une distance et des angles. Le système de coordonnées le plus couramment utilisé en mécanique newtonienne est celui des coordonnées rectangulaires ou cartésiennes. Il arrive cependant que l'on fasse usage de coordonnées polaires pour certains problèmes.⁶ La position d'un point matériel dans cet espace est donc entièrement déterminée par trois coordonnées.

Notons que l'on représente souvent un solide par un point matériel que l'on situe en son centre d'inertie ; c'est le cas lorsque l'on étudie par exemple le mouvement des planètes et de leurs satellites. Dans le cas d'un solide indéformable⁷ qui se meut librement dans un espace à trois dimensions, le nombre de coordonnées indépendantes, appelées aussi « degrés de liberté », nécessaires à la détermination de sa position et de son orientation est de six.⁸

On distingue, au sein de la mécanique, entre la cinématique, qui est la description du mouvement des corps, et la dynamique, qui est l'étude des causes de ce mouvement. Les notions de base de la cinématique sont celles de position, de vitesse et d'accélération. Elles sont habituellement représentées par des vecteurs. Soit s la courbe décrite par la particule dans son mouvement et \mathbf{r} le rayon vecteur⁹ pointant de l'origine vers la particule. \mathbf{r} est appelé « vecteur position » et ses coordonnées (x, y, z) sont les coordonnées cartésiennes de position de la particule. La vitesse étant

⁵D'autres grandeurs peuvent caractériser l'état d'un point matériel, comme la charge. Je ne m'en occuperai pas dans cet exposé.

⁶On utilise les coordonnées cylindriques pour les problèmes qui font intervenir une direction privilégiée, par exemple lors de l'étude du mouvement de rotation d'un système autour d'un axe. On utilise les coordonnées sphériques pour l'étude des cas où une propriété ne dépend que de la distance à un point ; il peut alors être commode de considérer ce point comme l'origine O d'un système de coordonnées sphériques.

⁷Un système matériel est dit « solide indéformable » ou « corps rigide » si les distances mutuelles des points matériels qui le constituent ne varient pas au cours du temps. Je parlerai simplement de « solide » dans la mesure où je ne m'intéresserai pas aux problèmes de déformation.

⁸Ce sont les trois coordonnées de position nécessaires à la détermination de la position d'un point dans un espace à trois dimensions, et les trois coordonnées angulaires permettant de déterminer l'orientation du solide. Ces trois coordonnées angulaires sont appelées « angles d'Euler ».

⁹J'adopte dans toute cette étude la convention orthographique qui consiste à représenter les quantités vectorielles par des caractères gras.

la dérivée de la position par rapport au temps¹⁰, le vecteur vitesse peut alors être défini par l'équation $\mathbf{v} = \frac{d\mathbf{r}}{dt}$ ou $\mathbf{v} = \dot{\mathbf{r}}$. L'accélération est la dérivée de la vitesse par rapport au temps, ou la dérivée seconde de la position : $\mathbf{a} = \frac{d^2\mathbf{r}}{dt^2} = \frac{d\mathbf{v}}{dt}$ ou $\mathbf{a} = \ddot{\mathbf{r}}$.¹¹

L'état d'un système à un instant t est caractérisé par les positions et les vitesses de chacune de ses particules à cet instant t , jointes à leurs caractéristiques intrinsèques, comme la masse. L'état d'un système à chaque instant est donc spécifié par deux ensembles de coordonnées, celles des positions \mathbf{r} et celles des quantités de mouvement \mathbf{p} (la *quantité de mouvement* ou *impulsion* d'un point matériel est le produit de sa masse par sa vitesse $m\mathbf{v}$) de chacune de ses particules. Si ces dernières se meuvent librement dans un espace à trois dimensions, la position et la quantité de mouvement de chacune d'elle aura trois composantes. L'état d'une particule unique est donc donné par six coordonnées, trois pour la position et trois pour la quantité de mouvement. L'état d'un système à n particules est donc entièrement spécifié par un total de $6n$ coordonnées.

1.2 Lois de la dynamique

La dynamique est l'étude des causes du mouvement. Le concept central de la dynamique newtonienne est celui de « force ». Les trois lois de Newton, également appelées « axiomes » du mouvement, gouvernent la dynamique des systèmes de particules.¹²

1.2.1 Première loi de Newton : le principe d'inertie

Le principe d'inertie stipule que le mouvement d'un point matériel auquel aucune force extérieure ne s'applique est rectiligne et uniforme ; c'est-à-dire que, sans l'action d'une force, la vitesse d'un point matériel ne peut changer, son mouvement ne peut subir aucune accélération. Ce principe donne la définition des référentiels¹³ galiléens

¹⁰La dérivée est calculée par le passage à la limite :

$$\frac{d\mathbf{r}}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{r}_2 - \mathbf{r}_1}{\Delta t}; \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t} = \frac{ds}{dt}$$

¹¹Le point ($\dot{\mathbf{r}}$) est utilisé pour symboliser la dérivée première par rapport au temps, et le double point ($\ddot{\mathbf{r}}$) pour symboliser la dérivée seconde par rapport au temps.

¹²Le statut de ces différents principes, et la question de savoir s'ils sont des postulats, des définitions ou des lois empiriques a fait l'objet de nombreux débats depuis le dix-huitième siècle. En particulier, la deuxième loi peut être interprétée comme un postulat fondamental ou comme une définition de la force et de la masse. Je les expose ici sans considérer ces questions, que l'on retrouvera partiellement au chapitre 3.

¹³Un référentiel est un système d'axes lié à un observateur auquel est associé un temps t .

(ou référentiels d'inertie) : ce sont les référentiels où ce principe, associé à un espace euclidien, est valable.¹⁴

Le principe de relativité galiléenne stipule que les lois de la mécanique sont invariantes relativement à une transformation galiléenne. Cela signifie que tous les repères galiléens sont équivalents : si deux observateurs sont animés d'un mouvement de translation uniforme l'un par rapport à l'autre, alors les mêmes lois du mouvement doivent s'appliquer à chacun d'eux.

1.2.2 Deuxième loi de Newton : le principe fondamental de la dynamique (PFD)

On peut définir la *force* comme la cause de l'accélération d'un point matériel. On représente la force, comme les grandeurs cinématiques, par un vecteur. La force \mathbf{F} agissant sur un point en modifie la vitesse, en fonction de la masse m du point.¹⁵

Le *principe fondamental de la dynamique* (ou *relation fondamentale de la dynamique*, parfois noté « PFD ») stipule que, dans un référentiel galiléen, la force qui s'exerce sur un point matériel est égale à la dérivée par rapport au temps du vecteur « quantité de mouvement », soit :

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \quad (1)$$

Cette équation peut aussi s'écrire

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}) \quad (2)$$

Quand la masse est constante, on peut écrire le PFD sous sa forme plus familière :

$$\mathbf{F} = m\mathbf{a} \quad (3)$$

1.2.3 Troisième loi de Newton : le principe des actions réciproques ou principe de l'égalité et de la réaction

À strictement parler, on quitte ici le domaine de l'étude du mouvement d'un point matériel isolé, pour aborder celle des systèmes de points matériels. Considérons deux points matériels i et j en interaction. Notons \mathbf{F}_{ij} la force exercée par le point i sur

¹⁴Ce principe peut être considéré comme une simple définition, comme un postulat théorique, ou comme l'hypothèse non triviale de l'existence de référentiels galiléens.

¹⁵La mécanique newtonienne ne distingue pas entre masse inertielle et masse gravitationnelle. L'inertie du corps, c'est-à-dire sa résistance au mouvement, est caractérisée par la quantité de mouvement $\mathbf{p} = m\mathbf{v}$.

le point j et \mathbf{F}_{ji} la force exercée par j sur i ; le principe des actions réciproques – ou principe de l'égalité de l'action et de la réaction – stipule que ces deux forces sont opposées, et égales en module, ce qui se traduit par la relation :

$$\mathbf{F}_{ij} = -\mathbf{F}_{ji} \quad (4)$$

Remarques

1. Ce qui précède concerne l'étude de mouvements rectilignes. Quand on sort de ce domaine, on emploie les équivalents angulaires de la position, de la vitesse et de l'accélération linéaires (voir l'annexe A.1)¹⁶.
2. Les trois lois exposées s'étendent sans modification profonde aux systèmes de particules (voir l'annexe A.2). Il faut cependant distinguer entre les *forces externes* agissant sur les points matériels et dues à des causes extérieures au système, et les *forces internes* agissant sur un point matériel i et dues aux autres points matériels du système. La deuxième loi de Newton prend alors cette forme :

$$\sum_j \mathbf{F}_{ji} + \mathbf{F}_i^{(e)} = \dot{\mathbf{p}}_i \quad (5)$$

où $\mathbf{F}_i^{(e)}$ représente une force externe et \mathbf{F}_{ji} est la force interne due à la particule j agissant sur la particule i .

3. Dans les présentations courantes de la formulation newtonienne, on ajoute généralement à l'exposé des trois lois du mouvement celui de certains *théorèmes de conservation*. Certains sont des conséquences directes des lois du mouvement ; d'autres nécessitent l'introduction de notions (celles d'énergie cinétique, d'énergie potentielle et de travail) qui ne figurent pas dans l'énoncé de ces lois, et qui sont par conséquent définies *ad hoc* (voir l'annexe A.3) ; ces notions sont en revanche les concepts fondamentaux des formulations analytiques de la mécanique.

1.3 En pratique

Un problème de mécanique consiste, typiquement, à prédire l'évolution dynamique d'un système, c'est-à-dire l'évolution de la position et de la vitesse de ses

¹⁶Les démonstrations trop longues pour figurer dans le texte principal, ainsi que certains développements importants pour une étude approfondie de la mécanique, mais qui ne servent pas directement mon propos, sont placés à la toute fin de ce travail, en annexes (à partir de la page 553). J'y renvoie régulièrement au cours de l'exposé qui suit.

différents points au cours du temps. La deuxième loi, ou principe fondamental de la dynamique, gouverne l'évolution dynamique de tout système : elle détermine à elle seule les trajectoires possibles pour ce système. En représentant l'état mécanique d'un système au moyen d'équations qui ont la forme du PFD, on est en principe en mesure d'effectuer des calculs qui permettent de prédire l'état de ce système à tout instant t . Cela signifie que la solution d'un problème de mécanique dans le cadre newtonien est apportée par la résolution d'une série d'équations différentielles de la forme suivante :

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = \mathbf{F}_i = \sum_j \mathbf{F}_{ij} + \mathbf{F}_i^{(e)} \quad (6)$$

En un sens, tout ce que la théorie newtonnienne dit à propos du mouvement des corps est contenu dans l'énoncé des trois lois du mouvement, qui offrent une représentation mathématique de tous les phénomènes du mouvement. Cependant, telle quelle, l'équation (6) est un principe abstrait : elle donne la forme générale des équations du mouvement, mais elle ne permet pas encore de dire quoique ce soit du comportement d'un système ou d'un type de systèmes. C'est, comme le note Kuhn (1970a, 1969), une « esquisse de loi ».¹⁷

L'utilisation de la théorie de Newton pour résoudre des problèmes de mécanique nécessite donc un apprentissage qui dépasse celui des trois lois de Newton. Prédire et expliquer des phénomènes particuliers implique d'être en mesure de les représenter au moyen de la deuxième loi de Newton, c'est-à-dire de donner un contenu à cette forme de loi de telle sorte qu'elle soit bien la représentation d'un système particulier.

Plusieurs étapes sont nécessaires pour obtenir la description de l'évolution mécanique d'un système particulier.¹⁸ Il faut tout d'abord *spécifier la nature des forces* en

¹⁷Voir aussi (Cartwright, 1999, p. 180) et (Giere, 1988). L'idée selon laquelle les lois théoriques ne « disent rien » par elles-mêmes est un des arguments courants contre les conceptions des théories scientifiques qui les définissent comme un ensemble de principes joints à l'ensemble de leurs conséquences déductives, que Cartwright (1999, chap. 8), appelle les conceptions « distributeur automatique » [*vending machine view of theory*]; la plus célèbre de ces conceptions est celle des empiristes logiques, à laquelle une partie du chapitre 5 est consacrée. Ses défenseurs prennent bien évidemment en compte la nécessité d'ajouter à ces principes des règles permettant de les appliquer aux phénomènes, mais ils décrivent ces règles d'une manière qui, selon leurs critiques – que j'endors –, ne rend pas compte des opérations nécessaires à la transformation des lois théoriques en des descriptions effectives des phénomènes, opérations qui requièrent une grande part de créativité de la part des agents. J'examinerai ces problèmes au chapitre 6.

¹⁸L'analyse détaillée des opérations dans lesquelles consiste chacune de ces étapes est un des aspects du projet sur la modélisation conduit par Nancy Cartwright (1999) et par Morgan et Morrison (1999). Elle s'accompagne de la thèse selon laquelle les modèles sont des outils autonomes, qui permettent d'opérer une médiation entre la théorie et les phénomènes sans toujours dériver de la théorie elle-même.

présence : les forces externes appliquées au système et les forces internes s'exerçant entre les différentes particules du système. Il existe plusieurs types de forces (forces de frottement, force de gravitation, force de rappel...). On appelle « loi de force » ou « fonction de force » la forme particulière que prend le \mathbf{F} du principe fondamental de la dynamique selon le type de force en jeu. Une fois ces forces identifiées, on peut *écrire les équations du mouvement* du système. Ce sont des équations vectorielles ; il y en a une pour chaque particule et pour chaque composante de la direction. Elles gouvernent entièrement l'évolution dynamique du système, *pour un état initial donné*.

L'étape suivante consiste à donner les conditions cinématiques initiales du système, c'est-à-dire les positions et quantités de mouvement initiales de chacune de ses particules. Intégrer deux fois¹⁹ chaque équation conduit alors à une solution unique pour chaque particule. Cette solution est une fonction vectorielle de la position de la particule en fonction du temps $\mathbf{r}(t)$, et permet par conséquent de déterminer la position et la vitesse de la particule à chaque instant.

Comprendre la mécanique newtonienne, c'est-à-dire savoir l'utiliser comme outil de représentation et d'inférence, implique donc l'acquisition d'un savoir-faire qui dépasse la connaissance des hypothèses théoriques *in abstracto*. En particulier, cela implique, face à une situation concrète plus ou moins complexe, de savoir identifier les forces en présence, externes et internes, c'est-à-dire de savoir représenter un système particulier *comme un système newtonien*. Selon les cas, la forme que prend alors le PFD peut le rendre méconnaissable aux yeux du novice. Comme Kuhn (1962/1970) l'a mis en évidence, les étudiants acquièrent typiquement ce savoir-faire au moyen des exercices proposés dans les manuels, qui leur présentent des exemples types, ou cas exemplaires [*exemplars*], dont la difficulté s'accroît au fur et à mesure de l'apprentissage, et qui les préparent à savoir trouver, par la suite, la solution à des problèmes jamais rencontrés auparavant.²⁰

1.3.1 Exemples types d'application des lois de Newton

La plupart des manuels de mécanique, en plus de l'énoncé des trois lois de Newton, considèrent comme partie intégrante de la théorie mécanique certaines lois de force comme celles de la gravitation universelle²¹ ou de certaines forces de rappel. Ces lois, que l'on peut appeler « lois phénoménales », sont des hypothèses générales, au sens où elles ne concernent pas un système en particulier, mais tous les systèmes soumis

¹⁹ Les équations du mouvement sont des différentielles du second ordre ; c'est la raison pour laquelle on a besoin de deux ensembles de valeurs initiales (positions et quantités de mouvement).

²⁰ La notion kuhnienne de cas exemplaire fera l'objet d'une analyse détaillée au chapitre 6.

²¹ Voir l'annexe A.4.

à un certain type de forces. Elles sont cependant moins abstraites que le PFD, dont elles sont une spécification : contrairement au PFD, elles ont un contenu empirique, au sens où elles disent quelque chose à propos d'un certain type de forces.

Comme l'a noté Giere (1988, p. 6), les manuels de mécanique sont la plupart du temps organisés en chapitres dont chacun est consacré à l'examen d'un type particulier de forces, et à la résolution de plusieurs problèmes exemplaires. Dans ce qui suit, j'expose deux de ces problèmes exemplaires, que les manuels les plus élémentaires ne manquent jamais de présenter. Les systèmes concernés sont des masses ponctuelles, ce qui évite le problème d'avoir à spécifier les forces internes au système (celles qui s'exercent entre les particules). Dans les deux cas, la force qui s'exerce sur le système est une force de rappel.

Le système masse-ressort. Loi de Hooke. La loi de Hooke gouverne la force de rappel s'exerçant par exemple dans un système (voir figure 1) constitué d'une masse m liée à un ressort accroché à un support fixe²² ; le ressort commence par s'étirer pour équilibrer la force de gravité ; une fois à l'équilibre, on peut étudier le déplacement vertical de la masse à partir de sa position d'équilibre.

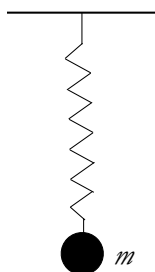


FIG. 1 – Système masse-ressort.

On appelle x le déplacement vertical et on suppose que l'allongement du ressort est parfaitement linéaire, auquel cas la force de rappel, lorsque le ressort est tiré, est

²²Le système masse-ressort est l'exemple le plus simple de ce que l'on appelle « oscillateur harmonique ». L'oscillateur harmonique est un exemple du mouvement périodique d'une grande importance car il sert de modèle dans beaucoup de problèmes de physique classique ou quantique. Il désigne le mouvement d'un point au voisinage d'une position d'équilibre stable. Les systèmes classiques qui sont des réalisations de l'oscillateur harmonique comprennent tous les systèmes stables quand on les écarte un peu de leur position d'équilibre. La notion d'oscillateur harmonique désigne donc le mouvement d'un point au voisinage d'une position d'équilibre stable. Les propriétés les plus importantes de l'oscillateur harmonique sont les suivantes : 1. la fréquence du mouvement est indépendante de l'amplitude de l'oscillation ; 2. on peut superposer linéairement les effets de plusieurs forces appliquées.

exactement proportionnelle au déplacement de la masse par rapport à sa position d'équilibre. La loi de Hooke stipule ainsi que la force exercée par un ressort est proportionnelle à son extension. La force est donc $-kx$ (le signe « $-$ » indique que c'est une force de rappel). Dans le cas du ressort, la constante k est interprétée comme une mesure de la raideur du ressort. Le principe fondamental de la dynamique s'écrit alors :

$$F = ma = \frac{md^2x}{dt^2} = -kx \quad (7)$$

où k est la constante de proportionnalité.

On peut alors intégrer l'équation une première fois pour obtenir $v(t)$, puis une seconde fois pour obtenir $x(t)$. Pour cela, la méthode la plus simple est de définir $\omega^2 = \frac{k}{m}$; l'équation 7 devient alors $a + \omega^2 x = 0$; les solutions pour la position et la vitesse en fonction du temps ont la même forme générale harmonique représentée par la fonction :

$$f(t) = A \cos \omega t + B \sin \omega t \quad (8)$$

où A et B sont des constantes à déterminer à l'aide des conditions initiales. Par exemple, si le ressort est tiré puis relâché sans vitesse initiale, les conditions initiales sont $x(t=0) = A$ et $v(t=0) = 0$. Dans ce cas, la solution de l'équation se réduit à :

$$x(t) = A \cos \omega t \quad v(t) = -A \sin \omega t \quad (9)$$

Les graphes des deux fonctions ont une forme sinusoïdale. Pour les conditions initiales $x(t=0) = A$ et $v(t=0) = 0$, on obtient les graphes de la figure 2.²³

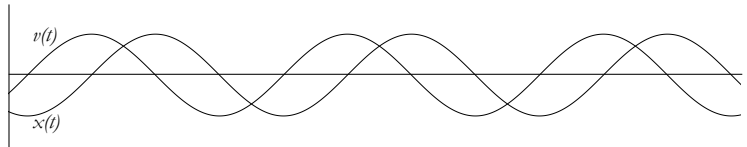


FIG. 2 – Position x et vitesse v de la masse m de la masse en fonction du temps t

²³Pour résoudre ce problème, on peut aussi utiliser la conservation de l'énergie. En effet, puisque la force est conservative (voir l'annexe A.3), $F(x) = -\nabla V(x)$ pour l'énergie potentielle $V(x)$. En intégrant $F(x) = ma$, on obtient l'équation $\frac{1}{2}mv^2 + V(x) = E$, avec E l'énergie totale. On résout pour $\frac{1}{v}$ et on intègre pour obtenir la solution.

Le pendule simple Un autre exemple de mouvement harmonique simple est offert par celui du pendule simple, dans la limite des petits angles d'oscillation. Le pendule simple (figure 3) consiste en une masse ponctuelle m à l'extrémité inférieure d'un fil inextensible sans masse de longueur l , sujette à un champ de gravitation uniforme $-mg$, et contrainte à se mouvoir le long d'un arc de cercle. On appelle θ son angle d'oscillation.

Afin de spécifier les forces s'exerçant sur le balancier du pendule, dont on suppose que c'est une masse ponctuelle, on décompose la force totale dans ses composantes. Choisissons la direction du fil comme axe des ordonnées et la tangente à l'arc de cercle décrit par le balancier à l'endroit du balancier comme axe des abscisses. La composante de la force dans la direction de l'accélération du balancier est $mg\sin\theta$. La longueur s de l'arc de cercle parcourue par le balancier à un certain instant, c'est-à-dire sa distance à l'origine, est donnée par $s = l\theta$. L'équation du mouvement du balancier devient alors :

$$-g\sin\theta = l\frac{d^2\theta}{dt^2} \quad (10)$$

Cette équation permet de vérifier que la période du pendule est proportionnelle à la racine carrée de sa longueur et indépendante de sa masse, comme l'avaient observé Galilée et Newton.

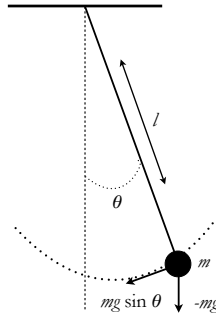


FIG. 3 – Pendule simple.

Notons que l'équation du mouvement n'a pas ici la forme simple de la loi de Hooke : l'expression de la force est dépendante du sinus de l'angle. Pour ce type de système, l'usage de coordonnées polaires est parfois plus avantageux que celui de coordonnées rectangulaires.²⁴ Sans que cela affecte les prédictions, la forme de la loi de Newton change avec un tel changement de coordonnées. Comme on va le voir, ce ne

²⁴On passe de coordonnées rectangulaires en coordonnées polaires par la transformation suivante : $x = r\cos\theta$ et $y = r\sin\theta$.

sera pas le cas avec les équations lagrangiennes et hamiltoniennes du mouvement.²⁵

Comme le montrent ces deux exemples simples – système masse-ressort et pendule simple –, l’application de la deuxième loi de Newton à la résolution de problèmes exige des compétences qui dépassent la seule capacité à opérer des déductions mathématiques au moyen d’équations différentielles : dans les deux cas, écrire la loi sous la forme adéquate exige d’être capable de représenter le problème – souvent en s’aidant d’un schéma – de façon à pouvoir identifier les forces pertinentes et à choisir les coordonnées les plus adaptées.

Remarque : la notion de modèle Les systèmes représentés dans les figures 1 et 3 sont des systèmes *idéaux* : la masse m est supposée ponctuelle, et le ressort comme le fil du pendule sont supposés ne pas avoir de masse ; il n’y a ni frottements, ni résistance de l’air, et le champ de gravitation est supposé uniforme. L’application de la loi de Hooke à un ressort réel nécessite des hypothèses simplificatrices ou idéalizations qui conduisent à représenter ce ressort *comme un ressort idéal*. On appelle couramment ces systèmes idéaux qui satisfont parfaitement les lois fournies par la théorie des *modèles*. Je reviendrai, au chapitre 5, ainsi qu’au chapitre 6 (section 2), sur la notion de modèle, qui fait l’objet de nombreux débats en philosophie des sciences, mais dont l’analyse m’éloignerait de mon propos dans le présent chapitre. Insistons cependant dès à présent sur la chose suivante : en tant qu’ils ont la forme d’un système concret (un ressort, ou, dans l’exemple suivant, un pendule), ces modèles permettent aux agents de se représenter ce que disent les lois de la mécanique au moyen d’un système (imaginaire) qui y obéirait parfaitement. Autrement dit, et pour employer un langage métaphorique, tout en étant imaginaires, ces modèles permettent de donner de la chair aux lois encore abstraites en se les *figurant*. En ce sens, ils sont des modèles *de la théorie*, qu’ils permettent d’illustrer.

En outre, dans la mesure où l’application des hypothèses de la théorie aux phénomènes peut être décrite comme une opération consistant à représenter, par exemple, un ressort réel *comme* un ressort idéal, ce dernier est aussi un modèle *du système*, au sens où il en est une représentation, qui porte en elle une hypothèse théorique. En apprenant à prédire le comportement d’un tel système idéal, on apprend à identifier et à représenter les forces en présence dans un système réel de telle sorte qu’il devient possible d’utiliser la loi de Hooke pour en prédire le comportement.²⁶

²⁵Pour plus de détail sur ce point, voir (North, 2009a, pp. 12-15).

²⁶Sur la notion de modèle, voir, entre autres, (Giere, 1988; Cartwright, 1999; Morgan et Morrison, 1999; Frigg, 2002, 2006; Suárez, 1999, 2003; Godfrey-Smith, 2006).

1.3.2 Limites pratiques de la formulation newtonienne

Les différentes opérations décrites ci-dessus (identification des forces en présence, écriture des équations, résolution des équations) s'avèrent parfois difficiles, voire impossible à effectuer. Pour certains systèmes, c'est l'étape de la mise en équation qui est pratiquement impossible ; dans d'autres cas, c'est la résolution des équations qui pose problème, car certaines équations ainsi obtenues n'ont pas de solution analytique (c'est-à-dire qu'elles ne sont pas intégrables). Dans ce qui suit, je présente un cas typique de système que la formulation newtonienne ne permet pas, en pratique, de décrire, et qui motive l'introduction de la formulation analytique.

Dans certains cas, l'utilisation du principe fondamental de la dynamique s'avère pratiquement impossible. Spécifier les forces s'exerçant entre les particules d'un système afin d'écrire les équations du mouvement correspondantes n'est pas une opération qui va de soi ; pour certains systèmes, on rencontre à cette étape des difficultés insurmontables. C'est le cas de la plupart des systèmes mécaniques soumis à ce que l'on appelle des *contraintes* ou *liaisons*.

Un système *contraint* est un système dont les particules ne sont pas *libres*. Il existe dans la plupart des systèmes physiques des liaisons entre particules qui les empêchent de se mouvoir librement indépendamment les unes des autres. Contrairement à ce qui est le cas pour un système de n particules libres, un système contraint n'a pas $6n$ degrés de liberté. L'exemple le plus simple d'un système contraint est celui d'un corps rigide, où les contraintes sur les particules sont telles que les distances entre elles restent inchangées. D'autres exemples peuvent être fournis, comme celui des boules d'un boulier contraintes à un mouvement uni-dimensionnel, celui des molécules de gaz à l'intérieur d'un récipient, contraintes à se mouvoir uniquement à l'intérieur de ce récipient, celui d'une particule placée à la surface d'un solide sphérique, contrainte à ne se mouvoir qu'à la surface ou dans la région extérieure à la sphère, ou encore celui de deux pendules liés et contraints à se déplacer dans un plan (figure 4).²⁷ On distingue différents types de contraintes (voir l'annexe A.5). Notons simplement que

²⁷Lanczos (1970) propose cette liste d'exemples de systèmes contraints et de leur nombre de degrés de liberté :

Un degré de liberté : un piston se mouvant de haut en bas. Un corps rigide en rotation autour d'un axe fixe.

Deux degrés de liberté : une particule se mouvant sur une surface donnée.

Trois degrés de liberté : une particule se mouvant dans l'espace. Un corps rigide en rotation autour d'un point fixe.

Quatre degrés de liberté : deux composantes d'une étoile double en révolution sur le même plan.

Cinq degrés de liberté : deux particules maintenues à une distance constante l'une de l'autre.

Six degrés de liberté : deux planètes en révolution autour d'un soleil fixe. Un corps rigide se mouvant librement dans l'espace.

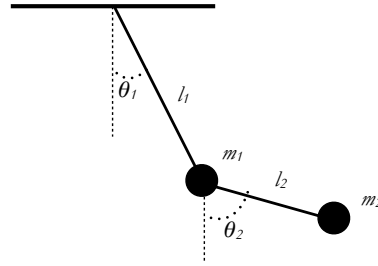


FIG. 4 – Deux pendules liés, contraints à se déplacer dans un plan.

les contraintes dites « holonomes », auxquelles je m'intéresserai presque exclusivement, peuvent être exprimées sous la forme d'équations reliant les coordonnées des particules et le temps ayant pour forme :

$$f(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, t) = 0 \quad (11)$$

Les contraintes introduisent *deux types de difficultés* dans la résolution des problèmes de mécanique selon l'approche newtonienne :

1. les coordonnées \mathbf{r}_i des particules ne sont pas indépendantes les unes des autres ; les équations du mouvement ne le sont par conséquent pas non plus ;
2. les forces de contrainte ne sont pas données. Elles figurent parmi les inconnues du problème. Comme le dit Herbert Goldstein (1950/2002, chap. 1, p. 13), « imposer des contraintes au système est simplement une autre manière de dire qu'il y a des forces présentes dans le problème qui ne peuvent être spécifiées directement mais dont on a connaissance par leur effet sur le mouvement du système ». Cela n'empêche pas d'écrire les équations du mouvement : on peut traiter les contraintes en laissant les forces de liaison indéterminées. Cependant, l'élimination des forces de liaison fait alors partie du problème de résolution des équations.

Comme on va le voir maintenant, la formulation lagrangienne de la mécanique permet de surmonter ces deux difficultés.

2 Mécanique analytique – Formulation lagrangienne

Les formulations analytiques permettent de traiter de nombreux phénomènes que la mécanique newtonienne ne peut pas traiter, comme, par exemple, le mouvement des fluides. En outre, la forme des équations lagrangiennes se retrouve en théorie des champs et en relativité (pour les premières) et en mécanique statistique et quantique

(pour les secondes).²⁸ Les exposés des principes de la mécanique analytique motivent souvent la présentation de cette approche des phénomènes du mouvement par le constat des limites d'applicabilité de la formulation newtonienne. Les outils mathématiques et les grands principes de la mécanique analytique peuvent être présentés comme une réponse à la double difficulté dégagée ci-dessus.

Les formulations lagrangienne et hamiltonienne sont toutes deux des formulations analytiques. Elles se distinguent entre autres, comme on va le voir, par le type de coordonnées qu'elles utilisent pour représenter l'état des systèmes. Je présente d'abord la formulation lagrangienne, dont plusieurs aspects lui sont communs à la formulation hamiltonienne. Je présenterai ensuite, beaucoup plus brièvement, la formulation hamiltonienne (section 3).

Dans ce qui suit, je présente d'abord (section 2.1) les outils mathématiques au moyen desquels les formulations analytiques représentent l'état des systèmes mécaniques. La différence entre ces outils et ceux qui sont mis en œuvre par la mécanique newtonienne est cruciale pour mon propos ; elle est au cœur de l'analyse de la section 4 du présent chapitre. Ensuite (sections 2.2 et 2.3), j'expose brièvement – sans entrer dans le détail des déductions, que je donne en annexe – les différents principes (principe de d'Alembert et principe de Hamilton) qui permettent de déduire les équations lagrangiennes du mouvement. On verra que ces principes, déductibles des principes newtoniens, mettent en jeu des concepts différents (énergies cinétique et potentielle, travail)²⁹ de celui qui est au fondement de la mécanique newtonienne (force). Enfin (section 2.4), je présenterai la manière dont les équations lagrangiennes du mouvement peuvent être utilisées pour résoudre des problèmes mécaniques.

2.1 Outils mathématiques de la représentation des systèmes mécaniques : espace de configuration et coordonnées généralisées

Les outils mathématiques utilisés pour représenter les systèmes en mécanique analytique peuvent être décrits comme une manière de répondre à la *première difficulté* énoncée à la fin de la section 1.3.2. La clef de la solution réside dans une des idées fondamentales de l'approche analytique, celle d'un espace de configuration à n dimensions dans lequel la configuration du système entier est représentée par un point, dont les coordonnées sont appelées « coordonnées généralisées » et sont au nombre

²⁸Présenter cela en détail m'éloignerait du cœur de mon propos, et dépasse mes compétences. Mon exposé vise uniquement à mettre en évidence les différences entre ces formulations pour des cas simples.

²⁹Historiquement, ces concepts ont été introduits après la formulation des principes de la mécanique analytique. Cependant, dans sa formulation contemporaine, le principe de Hamilton met en jeu ces concepts. Pour une histoire du concept d'énergie, voir (Harman, 1982b) et (Darrigol, 2001).

de n , correspondant au nombre de degrés de liberté des particules du système. Les coordonnées généralisées d'un espace de configuration n'ont pas forcément, comme on va le voir, la dimension de positions.

Considérons un système mécanique de N particules. Si ces particules sont libres de contrainte, le système a $3N$ coordonnées indépendantes ou degrés de liberté. Dans ce cas, l'espace de configuration coïncide avec l'espace tridimensionnel où la position de chaque particule est déterminée par 3 coordonnées. S'il existe des contraintes holonomes, exprimées par k équations de la forme de l'équation (16a), on peut éliminer à l'aide de ces équations k des $3N$ coordonnées. Il est alors possible de caractériser la configuration du système par $n = 3N - k$ variables indépendantes q_1, q_2, \dots, q_N , appelées « *coordonnées généralisées*³⁰ », en fonction desquelles les coordonnées rectangulaires de position $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ sont exprimées par des équations qui contiennent *implicitement* les contraintes. Cela signifie que, contrairement à ce qui est le cas dans le cadre de la formulation newtonienne, qui représenterait ce système avec les $3N$ coordonnées de position et devrait spécifier *explicitement* les équations des forces qui maintiennent les contraintes entre les particules (ou alors se trouverait confrontée au problème de devoir éliminer les forces de liaison laissées indéterminées pour résoudre les équations), on n'a pas besoin de connaître ces forces pour représenter la configuration du système.³¹

Les transformations de coordonnées³² conduisant aux n coordonnées généralisées du système ont la forme suivante :

$$x_1 = f_1(q_1, \dots, q_{3N}), \quad (12)$$

³⁰Il y a une ambiguïté dans l'usage de l'expression « coordonnées généralisées » : au sens strict, elle désigne ce que l'on pourrait appeler les « positions généralisées » ; en ce sens, les coordonnées généralisées sont au nombre de n pour un système à n degrés de liberté. Parfois, l'expression désigne, en un sens plus large, l'ensemble formé par les positions et les vitesses généralisées ; en ce dernier sens, elles sont alors au nombre de $2n$ pour un système à n degrés de liberté.

³¹Si la contrainte est non holonome (voir l'annexe A.5), les équations qui l'expriment ne peuvent être utilisées pour éliminer les coordonnées dépendantes. L'exemple d'un objet roulant sur une surface sans glissement le montre : la condition de roulement est une condition différentielle qui ne peut être donnée sous forme intégrée qu'après résolution du problème. On peut toujours, si la contrainte est non intégrable, introduire les équations différentielles de liaison dans le problème en même temps que les équations différentielles du mouvement et éliminer les équations dépendantes, par la méthode dite des « multiplicateurs de Lagrange » (voir Lanczos, 1970, pp. 43-48 et, pour une analyse de l'interprétation physique de cette méthode, Butterfield, 2004, pp. 46-49). Cependant, je supposerai presque toujours dans mon exposé que les contraintes sont holonomes ; cette restriction est sans effet sur mon propos.

³²La représentation de l'état – ou configuration – d'un système à l'aide de coordonnées généralisées est une généralisation de l'idée de changement de coordonnées (comme dans le cas du passage de coordonnées rectangulaires x, y, z , à des coordonnées polaires r, θ, ϕ).

$$\begin{aligned}
 \dots &= \dots \\
 \dots &= \dots \\
 z_N &= f_{3N}(q_1, \dots, q_{3N}).
 \end{aligned}$$

De même que l'on associe aux trois nombres x , y et z un point dans un espace tri-dimensionnel, on associe aux n nombres q_1, q_2, \dots, q_n un point P dans un espace à n dimensions. De même, si l'on peut associer une courbe et le mouvement d'un point le long de cette courbe aux équations

$$\begin{aligned}
 x &= f(t) \\
 y &= g(t) \\
 z &= h(t)
 \end{aligned} \tag{13}$$

on considère que les équations correspondantes

$$\begin{aligned}
 q_1 &= q_1(t) \\
 \dots &= \dots \\
 \dots &= \dots \\
 q_n &= q_n(t)
 \end{aligned} \tag{14}$$

représentent la solution d'un problème dynamique : leurs solutions donnent les coordonnées d'un point P dans un espace à n dimensions en mouvement le long d'une courbe donnée de cet espace, qui représente l'état du système étudié. Par exemple, la position d'un corps rigide est symbolisée par un seul point dans un espace à six dimensions. Les diverses positions du corps sont représentées par des points de cet espace.

Une des caractéristiques de la mécanique analytique est de ne pas spécifier le type de coordonnées utilisé pour représenter le mouvement d'un système : n'importe quel ensemble de paramètres qui permet de caractériser la position d'un système mécanique peut être choisi comme système de coordonnées. Comme on va le voir, cela a pour conséquence importante le fait que les équations de la mécanique analytique sont indépendantes du système de coordonnées dans lesquelles elles sont formulées. Cela permet, entre autres choses, de relier la dynamique avec des géométries d'espace à plusieurs dimensions.

2.2 Principe de d'Alembert et équations de Lagrange

Rappelons à présent la seconde difficulté présentée par les systèmes contraints (voir page 50) : les forces de contrainte font partie des inconnues du problème. Il faut

donc trouver un moyen de formuler la mécanique qui autorise à se passer des forces de contrainte. Le *principe du travail virtuel*, généralisé de la statique à la dynamique par le *principe de d'Alembert*, rend cela possible. On va voir que les équations du mouvement de Lagrange, équations fondamentales de l'approche analytique de la mécanique, en sont déductibles. Le principe de d'Alembert n'est pas en tant que tel un principe variationnel (comme les différentes formes du principe de moindre action, que je présenterai ensuite), c'est-à-dire qu'il n'implique pas la minimisation d'une intégrale, mais on verra que les différents principes fondamentaux de l'approche analytique de la mécanique sont déductibles les uns des autres.

2.2.1 Principe du travail virtuel

Le principe du travail virtuel³³ est un principe de la statique : il énonce une condition pour l'équilibre des systèmes contraints. Le *travail* W d'une force \mathbf{F} s'appliquant à une particule en mouvement du point 1 au point 2 est défini par $W_{12} = \int_1^2 \mathbf{F} \cdot d\mathbf{s}$. Appelons *déplacement virtuel* $\delta \mathbf{r}_i$ d'un point i tout changement infinitésimal de la position de ce point compatible avec les contraintes.³⁴ Le principe du travail virtuel affirme qu'un système est en équilibre si et seulement si, pour tout déplacement virtuel $\delta \mathbf{r}_i$ du point \mathbf{r}_i , la somme des travaux virtuels des forces qui s'y appliquent (\mathbf{F}_i) est nulle.

$$\sum_i \mathbf{F}_i \cdot \delta \mathbf{r}_i = 0 \quad (15)$$

Les forces dont il s'agit n'incluent pas les forces de liaison : le principe du travail virtuel énonce une condition pour l'équilibre d'un système contraint en se passant des forces qui maintiennent les contraintes. En tant que tel, il repose sur une *restriction* à ce que Jeremy Butterfield (2004, p. 44) appelle des contraintes « idéales », c'est-à-dire des contraintes telles que les forces qui les maintiennent ne produiraient aucun travail lors d'un déplacement virtuel. Les systèmes contraints – ou systèmes à liaison – sont définis comme des systèmes pour lesquels tout déplacement virtuel est réversible et dont une force arbitrairement faible appliquée dans la direction d'un déplacement possible suffit à briser l'équilibre (Darrigol, 2007, p. 763). Cette restriction exclut, entre autres, les forces de frottement.³⁵

³³Ce principe est dû à Jean Bernoulli et d'Alembert (1743), qui ont élargi les principes de l'équilibre énoncés, entre autres, par Stévin, aux systèmes à liaisons. Pour un exposé plus complet de ce principe, voir (Lanczos, 1970, pp. 75-78) et (Darrigol, 2007, pp. 763-766).

³⁴La notation δ pour indiquer une variation (d'une quantité quelconque, et non pas seulement de la position) *possible* ou *virtuelle*, par opposition à une variation actuelle notée d , a été introduite par Lagrange.

³⁵Lanczos (1970, pp. 76-77) présente cette restriction comme le postulat fondamental de la mé-

2.2.2 Principe de d'Alembert

Le principe du travail virtuel satisfait nos besoins en ce qu'il ne contient pas les forces de liaison, mais il n'est valable qu'en statique. Le principe de d'Alembert est une extension de ce principe à la dynamique. Comme on va le voir, les équations lagrangiennes du mouvement en sont déductibles.

Le procédé qui permet le passage de la statique à la dynamique ou, si l'on veut, la réduction de la dynamique à la statique, consiste à traiter l'opposé du produit $m\mathbf{a}$ comme une force. On réécrit alors le principe fondamental de la dynamique ($\mathbf{F} = m\mathbf{a}$) sous la forme :

$$\mathbf{F} - m\mathbf{a} = 0 \quad (16)$$

On sait que l'annulation d'une force dans la mécanique newtonienne implique l'équilibre. L'équation (16) affirme donc que l'addition de la force $-m\mathbf{a}$, appelée « force d'inertie », aux autres forces agissantes produit l'équilibre. Cela signifie que tout système de forces est en équilibre si on ajoute aux forces appliquées les forces d'inertie.³⁶ Si l'on est en possession d'un critère pour l'équilibre d'un système mécanique, on peut alors étendre ce critère à un système en mouvement. Il suffit en effet d'ajouter la force d'inertie aux forces en présence.

Combiné au principe du travail virtuel, le principe de d'Alembert permet d'affirmer que *la somme des travaux virtuels des forces appliquées et des forces d'inertie s'annule pour tout déplacement compatible avec les liaisons $\delta\mathbf{r}_i$* :

$$\sum (\mathbf{F}_i - m_i\mathbf{a}_i) \cdot \delta\mathbf{r}_i = 0 \quad (17)$$

Insistons sur le fait que les forces de liaison ne figurent pas dans ce principe (les \mathbf{F}_i ne comprennent pas les forces de liaison). Dans les cas où la trajectoire virtuelle coïncide avec la trajectoire réelle, le principe de d'Alembert permet de déduire le théorème de conservation de l'énergie.³⁷ Pour mon propos, l'aspect le plus important du principe de d'Alembert est qu'il permet de déduire les équations lagrangiennes du mouvement.

canique analytique et le baptise « postulat A ». En effet, dès lors qu'il est étendu aux systèmes dynamiques par le principe de d'Alembert, il permet de déduire les principes variationnels fondamentaux de la mécanique analytique. Je reviendrai sur ce point au chapitre 3, section 4.4 (page 149).

³⁶Pour une analyse des hypothèses qui sous-tendent le principe de d'Alembert, voir (Darrigol, 2007, pp. 771-773).

³⁷Voir l'annexe A.6 et, pour plus de détails, (Lanczos, 1970, pp. 94-96). Pour l'examen détaillé des conséquences du principe de d'Alembert, voir (Lanczos, 1970, pp. 94-110).

2.2.3 Les équations de Lagrange. Les concepts d'énergies cinétique et potentielle et de travail

En mécanique lagrangienne, la dynamique des systèmes est entièrement gouvernée par un ensemble d'équations, appelées « équations de Lagrange » ou « équations lagrangiennes ». Elles sont donc les analogues, dans cette formulation de la mécanique, des équations newtoniennes, dont on a vu qu'elles avaient la forme du principe fondamental de la dynamique.

Les équations lagrangiennes, dont on verra plus loin qu'elles sont déductibles du principe de Hamilton – lui-même déductible du principe de d'Alembert – sont des conséquences du principe de d'Alembert. Leur déduction repose sur l'expression du principe de d'Alembert (équation 17) à l'aide de *coordonnées généralisées* et sur la restriction à des systèmes holonomes. Une chaîne de déductions mathématiques³⁸ conduit aux équations suivantes :

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} = Q_i \quad (18)$$

où T est l'*énergie cinétique du système*, définie par $T = \frac{1}{2}m\mathbf{v}^2$, et où Q_i représente les composantes d'un vecteur de l'espace de configuration à n dimensions, que l'on appelle la « force généralisée », et qui suffit à représenter l'action dynamique de toutes les forces appliquées au système (de même que le mouvement du système peut être représenté comme le mouvement d'une particule unique dans cet espace à n dimensions).

Dans le cas des forces que l'on appelle « conservatives », le travail total effectué par la force Q_i au cours d'un mouvement ne dépend que de la position des points extrêmes et non de la trajectoire particulière effectivement suivie.³⁹ Il devient alors possible d'exprimer le travail W comme la variation d'une quantité qui dépend seulement de la position des points terminaux. Cette quantité peut être notée $-V$, de telle sorte que l'on a la relation $Q_i = -\frac{\partial V}{\partial q_i}$. V désigne l'*énergie potentielle* du système dans le champ de force Q_i .⁴⁰

³⁸La déduction, exposée en détail par Butterfield (2004, pp. 51-53) et par Goldstein (1950/2002, pp. 16-21), est reproduite en annexe (A.7).

³⁹Cela signifie que le travail W effectué autour d'une orbite fermée est nul, c'est-à-dire que $\oint \mathbf{F} \cdot d\mathbf{s} = 0$. Le travail doit alors être indépendant du chemin suivi pour intégrer $\int_1^2 \mathbf{F} \cdot d\mathbf{s}$ entre les points 1 et 2.

⁴⁰Les systèmes conservatifs sont ceux qui satisfont le théorème de conservation de l'énergie selon lequel l'énergie totale de la particule $T + V$ est conservée au cours du mouvement. En effet, pour une masse constante, $W_{12} = T_1 - T_2$ (c'est le théorème de l'énergie cinétique). Comme, dans le cas des systèmes conservatifs, on a aussi $W_{12} = V_1 - V_2$, on trouve bien $T_1 + V_1 = T_2 + V_2$. Signalons cependant qu'il existe des forces dérivables d'une fonction scalaire qui ne sont pas conservatives :

On définit alors une fonction L , appelée *lagrangien*, telle que $L = T - V$, qui permet de reformuler les équations (35a) comme suit :

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad (19)$$

C'est la forme standard des *équations lagrangiennes du mouvement*, qui gouvernent la dynamique des systèmes mécaniques.

Les équations lagrangiennes du mouvement, contrairement aux équations newtoniennes, n'emploient pas le concept de force, représenté en mécanique newtonienne par une quantité vectorielle qui détermine le mouvement des systèmes. La quantité fondamentale des équations lagrangiennes, le lagrangien L , est une fonction scalaire, et non plus vectorielle, qui est définie en termes d'énergie.⁴¹

Les concepts d'énergies cinétique et potentielle et de travail, dont j'ai signalé (page 42) qu'ils étaient la plupart du temps introduits par les manuels qui s'en tiennent à la formulation newtonienne par des définitions *ad hoc* permettant d'exposer le théorème de conservation de l'énergie, deviennent donc les concepts fondamentaux au moyen desquels la mécanique lagrangienne représente les phénomènes du mouvement : la dynamique des systèmes y est entièrement gouvernée par une fonction définie en termes d'énergie. Dans ce cadre, la notion de force devient une quantité secondaire, dérivée de celle, plus centrale, de travail.⁴²

Insistons sur un point crucial : les concepts énergétiques peuvent être définis au moyen du concept de force et, inversement, le concept de force peut être dérivé de

c'est le cas des forces dérivant d'un potentiel U *dépendant du temps*, et qui n'est alors pas équivalent à l'énergie potentielle. Dans la terminologie adoptée par Lanczos (1970, p. 30), les forces dérivant d'un potentiel, qu'elles soient conservatives ou non, sont appelées *monogéniques* ; les forces qui ne dérivent pas d'un potentiel (comme par exemple les forces de frottement), sont dites *polygéniques*. Dans les cas les plus généraux de forces monogéniques non conservatives, les deux scalaires fondamentaux de la mécanique ne sont pas T et V , mais T et U . Dans le cas particulier des forces conservatives, $V = -U$. Je m'intéresserai presque exclusivement, dans tout mon exposé, au cas des forces conservatives.

⁴¹Le concept d'énergie a été introduit après la formulation des équations du mouvement par Lagrange. Cependant, dans la formulation contemporaine des équations de Lagrange, le lagrangien est défini en termes d'énergie. Comme j'y ai déjà insisté, ce chapitre est consacré aux formulations de la mécanique telles qu'elles sont enseignées aujourd'hui.

⁴²Comme le suggère Jeremy Butterfield (2004, p. 42), on peut présenter les notions d'énergie cinétique et de travail comme les « analogues » dans le cadre de la représentation du mouvement des corps à l'aide d'un espace de configuration tel que défini ci-dessus, des expressions placées des deux côtés de l'équation exprimant la deuxième loi de Newton. L'inertie d'un corps représentée par l'expression $m\mathbf{a}$ et la force \mathbf{F} qui s'exerce sur lui sont « déstituées de leur rôle de concepts centraux par leurs cousines : respectivement, l'énergie cinétique et le travail fourni par ces forces » (Butterfield, 2004, p. 42). Pour une analyse plus détaillée du statut de ces concepts en mécanique analytique, voir l'annexe A.8.

celui de travail. La *définition* de ces différents concepts est strictement la même dans les formulations newtonienne et lagrangienne. Il n'y a pas, comme dans le cas du passage de la mécanique newtonienne à la mécanique relativiste, de modification dans la définition des concepts. En ce sens, les formulations newtonienne et lagrangienne semblent bien exprimer une seule et même théorie, dont on peut considérer qu'elles l'abordent sous des perspectives différentes qui conduisent à une organisation différente de son architecture conceptuelle. La seule différence qui apparaît d'emblée entre elles réside dans le fait que la méthode lagrangienne ne fait intervenir, comme on l'a vu, que les paramètres individuellement contrôlables du système (en ignorant les degrés de liberté interne).

Ce faisant, ces deux formulations offrent cependant une représentation très différente des phénomènes du mouvement. Le passage d'une mécanique de la force à une mécanique de l'énergie implique, entre autres, le passage d'une représentation *instantanée* et *locale* à une représentation *globale* du mouvement (les concepts d'énergies cinétique et potentielle et de travail concernent en effet la trajectoire globale des systèmes). Ce changement, qui s'accompagne, comme on l'a vu, d'un changement dans les outils mathématiques utilisés pour représenter les phénomènes du mouvement, a des conséquences, comme on le verra bientôt, sur la manière dont, en pratique, les équations de Lagrange sont utilisées pour résoudre des problèmes.

2.3 Les principes variationnels. Le principe de Hamilton

Le principe de d'Alembert est un principe *différentiel* : il part de la considération de l'état *instantané* d'un système et des déplacements virtuels infinitésimaux autour de cet état instantané. L'utilisation d'un espace de configuration et de coordonnées généralisées et la reformulation des principes de base et des équations qui l'accompagnent placent ce que je viens d'exposer dans le domaine de la mécanique analytique, historiquement mise en place par Lagrange (1788). Cependant, la méthode *variationnelle* proprement dite repose sur l'utilisation de principes *variationnels* ou *intégraux*, qui impliquent la minimisation d'une intégrale. Ceux qui sont utilisés en mécanique analytique⁴³ peuvent être considérés comme des reformulations du principe de d'Alembert.

Les principes variationnels d'Euler, Lagrange, Hamilton et Jacobi peuvent également être appelés « principes de minimum » ou « principes de moindre action », selon la manière dont on définit l'*action* (elle a toujours, en mécanique analytique, la dimension d'une énergie), et consistent à considérer le *mouvement total* d'un système

⁴³On trouve des principes variationnels dans d'autres domaines de la physique, comme par exemple en optique, avec le principe de Fermat.

entre les instants t_1 et t_2 . À partir des faibles variations virtuelles du mouvement total autour du mouvement réel dans l'espace de configuration, on peut déduire les équations lagrangienne du mouvement.⁴⁴ Les problèmes de mécanique analytique, qui peut alors être dite « variationnelle », consistent à *définir une intégrale qu'il s'agit de minimiser*.⁴⁵ Ainsi, l'approche variationnelle de la mécanique nécessite que l'on trouve la valeur stationnaire d'une intégrale donnée, et non pas seulement, comme dans le cas du calcul différentiel classique, de celle d'une fonction.⁴⁶

Le principe de Hamilton est la transformation la plus naturelle et la plus directe du principe de d'Alembert en un principe de minimum ; à partir de là, on peut obtenir par une spécification les formes plus anciennes du principe utilisées par Lagrange et Euler⁴⁷, ainsi que le principe de Jacobi (voir l'annexe A.12).

2.3.1 Dédution du principe de Hamilton à partir du principe de d'Alembert

Le principe de d'Alembert (équation 17) utilise une équation différentielle non intégrable : en effet, si le travail virtuel des forces appliquées peut être décrit comme une variation de l'énergie potentielle, le travail virtuel des forces d'inertie est une forme différentielle non réductible à la variation d'une fonction scalaire ; il doit être établi pour chaque particule séparément. Cela rend le principe de d'Alembert pratiquement inutilisable dans de nombreux cas. La transformation du principe de d'Alembert en un principe variationnel repose sur une intégration par rapport au temps, qui permet de donner une forme monogénique⁴⁸ au travail effectué par les forces d'inertie, en les

⁴⁴Pour le détail des procédés du calcul variationnel, voir (Lanczos, 1970, chap. 2).

⁴⁵On parlera, à propos de la mécanique variationnelle, de minimisation ou d'extrémisation d'intégrales données. En réalité, il suffit, pour les problèmes mécaniques, de trouver des valeurs stationnaires de ces intégrales, et non de réels *minima* ou *maxima*, en dépit de ce que suggère l'emploi des termes « *minimum* » ou « *extremum* ».

⁴⁶Un problème typique de minimisation d'intégrale, proposé et résolu par Jean Bernouilli (1696), est celui de la « brachistochrone » (la courbe de la plus rapide descente). Euler et Lagrange ont contribué aux progrès du calcul variationnel en formulant les équations qui permettent de trouver la valeur stationnaire d'une intégrale donnée. C'est Lagrange qui le premier montre que la minimisation d'une intégrale donnée exige des outils spécifiques, différents de ceux du calcul différentiel ordinaire, et qui donne vraiment naissance au calcul variationnel.

⁴⁷Pour une présentation du principe de moindre action dans la version qu'en donnent Euler et Lagrange, voir (Butterfield, 2004, pp. 63-68). Le nom de « principe de Hamilton », donné par Jacobi, a été adopté au cours du vingtième siècle.

⁴⁸Une force monogénique est une force dérivant d'un potentiel.

représentant par l'énergie cinétique T . On obtient alors⁴⁹ :

$$\delta \int_{t_1}^{t_2} L dt = \delta A = 0 \quad (20)$$

avec $L = T - V$ ⁵⁰, et $A = \int_{t_1}^{t_2} L dt$.

C'est le *principe de Hamilton*, qui établit que *le mouvement d'un système mécanique est tel que l'intégrale A , appelée « action » et définie comme la différence entre l'énergie cinétique T et l'énergie potentielle V , devient stationnaire pour toute variation arbitraire possible de la configuration du système, pourvu que les configurations initiale et finale soient données*.⁵¹

Notons que le raisonnement qui conduit du principe de d'Alembert au principe de Hamilton peut être effectué en sens inverse, en partant du postulat que δA s'annule pour des variations arbitraires de position. Les principes de d'Alembert et de Hamilton sont donc *équivalents* dans les cas des systèmes holonomes où les forces appliquées dérivent d'un potentiel. En revanche, pour les forces ne dérivant pas d'un potentiel, la transformation du principe de d'Alembert en un principe de minimum n'est pas possible. Le principe de d'Alembert est plus général, puisqu'il n'exige pas d'intégration par rapport au temps, et qu'il s'applique en principe aux systèmes holonomes et non holonomes, alors que le principe de Hamilton est restreint aux systèmes holonomes. Cependant, alors que le principe de d'Alembert représente le mouvement de façon instantanée (c'est un principe différentiel), le principe de Hamilton considère le mouvement de façon globale, ce qui le rend utilisable dans les cas où le principe de d'Alembert n'est pas intégrable.

Avec le principe de Hamilton, le problème de la dynamique est alors réduit à la recherche de cette intégrale scalaire A . La condition pour que cette intégrale soit stationnaire donne toutes les équations du mouvement.

2.3.2 Equivalence du principe de Hamilton et des équations de Lagrange

Le principe de Hamilton est une condition nécessaire des équations de Lagrange, que l'on a précédemment déduites du principe de d'Alembert ; pour une démonstration, voir (Whittaker, 1959, section 99, pp. 245-247) et (Lanczos, 1970, pp. 58-59). Il en est aussi une condition suffisante : on peut en effet déduire les équations de Lagrange à partir du principe de Hamilton. Cette déduction repose sur la représentation du mouvement d'un système dans un espace de configuration à n dimensions,

⁴⁹Pour l'exposé de la déduction complète, voir (Lanczos, 1970, pp. 112-113), et l'annexe A.9.

⁵⁰On suppose que, en plus de dériver d'un potentiel, la force est conservative (voir note 40).

⁵¹On peut dériver le théorème de conservation de l'énergie du principe de Hamilton, ce qui permet de mettre en évidence la relation entre l'énergie totale d'un système et le lagrangien.

où l'état du système à chaque instant est représenté par un unique point. La courbe décrite par ce point entre deux instants t_1 et t_2 est appelée « trajectoire de l'état du système ».

Le principe de Hamilton affirme que le mouvement d'un système conservatif holonome⁵², entre t_1 et t_2 , est tel que la variation de l'intégrale de ligne A est nulle :

$$\delta A = \delta \int_{t_1}^{t_2} L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) dt = 0 \quad (21)$$

La méthode variationnelle permet de déduire les équations de Lagrange du principe

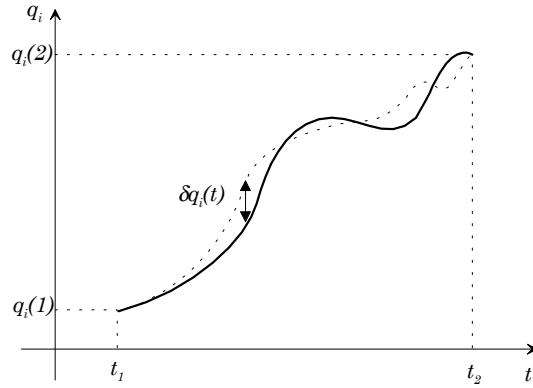


FIG. 5 – Trajectoire réelle et trajectoire variée de l'état d'un système. Image tirée du cours de Jean-Michel Raimond (2000).

de Hamilton. Elle consiste à considérer deux trajectoires possibles entre q_1 et q_2 (voir figure 5). L'une, notée $q_i(t)$, est la trajectoire effectivement suivie. L'autre, qu'on appelle « trajectoire variée », infiniment proche, correspond à chaque instant aux positions $q_i(t) + \delta q_i(t)$, où $\delta q_i(t)$ est un accroissement infinitésimal de la position. Ces deux trajectoires doivent obéir aux mêmes conditions initiales et finales. La condition selon laquelle l'action A est extrémale sur la trajectoire effectivement suivie $q_i(t)$ permet de déduire les équations lagrangiennes du mouvement (pour le détail de la déduction, voir l'annexe A.10).

2.4 En pratique

Dans la formulation lagrangienne de la mécanique, c'est une fonction scalaire, le Lagrangien L , qui gouverne l'évolution dynamique du système au cours du temps.

⁵²Le principe s'étend aussi aux systèmes monogéniques non conservatifs, ainsi qu'à certains systèmes qui ne sont pas à strictement parler monogéniques (voir Butterfield, 2004, pp. 69-70). Dans cet exposé, je me restreins aux systèmes conservatifs.

Pour les systèmes où l'on peut définir un lagrangien⁵³, on dispose d'un procédé très puissant pour établir les n équations du second ordre qui déterminent le mouvement d'un système à n degrés de liberté dans un espace à trois dimensions. Ici aussi, $2n$ valeurs initiales indépendantes sont nécessaires pour résoudre ces équations et connaître l'état du système à tout instant : en effet, le lagrangien $L(q, \dot{q})$ est une fonction des positions et des vitesses généralisées. Cependant, à la différence de ce qui est le cas dans le cadre de la formulation newtonienne, ces deux ensembles de coordonnées sont des coordonnées généralisées. De ce fait, les équations du mouvement ne changent pas de forme avec un changement de coordonnées, et cette formulation donne une grande liberté dans le choix des variables.

Avec les équations newtoniennes du mouvement, il faut opérer avec de nombreuses forces et accélérations *vectorielles*. Avec la méthode lagrangienne, il suffit de mettre en œuvre deux fonctions *scalaires*, T et V , ce qui simplifie énormément le problème. On peut à partir de là établir un procédé standard direct pour tous les problèmes de mécanique auxquels la formulation lagrangienne s'applique. Il suffit d'écrire T et V en coordonnées généralisées, de définir L et de les porter dans les équations de Lagrange pour obtenir les équations du mouvement. La transformation de T et V de coordonnées rectangulaires en coordonnées généralisées s'obtient par application des équations de transformation.

Une solution, trouvée par intégration des équations, a la forme d'une fonction ou trajectoire dans l'espace de configuration, $q(t)$. La résolution de chaque équation nécessite deux intégrations aboutissant au total à $2n$ constantes d'intégration. Dans un problème particulier, ces constantes sont déterminées par les conditions initiales, à savoir les valeurs initiales des n coordonnées généralisées q_j et des n vitesses généralisées \dot{q}_j .

2.4.1 Exemples de résolutions de problèmes

Particule en mouvement libre : équivalence des formulations newtonienne et lagrangienne Montrons d'abord, par un exemple simple, l'équivalence des résultats obtenus par les formulations newtonienne et lagrangienne. Prenons le cas d'une particule unique en mouvement libre. Les forces sont \mathbf{F}_x , \mathbf{F}_y et \mathbf{F}_z . Alors

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad (22)$$

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0 \quad (23)$$

⁵³C'est-à-dire pour des systèmes holonomes avec des forces appliquées dérivant d'un potentiel ordinaire ou généralisé.

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x}, \quad \frac{\partial T}{\partial \dot{y}} = m\dot{y}, \quad \frac{\partial T}{\partial \dot{z}} = m\dot{z}, \quad (24)$$

et les équations du mouvement sont

$$\frac{d}{dt}(m\dot{x}) = F_x, \quad \frac{d}{dt}(m\dot{y}) = F_y, \quad \frac{d}{dt}(m\dot{z}) = F_z \quad (25)$$

On a ainsi « retrouvé » les équations du mouvement dans leur forme newtonienne et montré l'équivalence des deux formulations pour traiter cet exemple.⁵⁴

Particule unique en mouvement le long d'une ligne droite Toujours sur la base d'exemples simples, que la formulation newtonienne permet de résoudre, je voudrais mettre en évidence la différence des procédés mis en œuvre par les deux formulations pour résoudre un problème.

Prenons d'abord le cas d'une particule unique en mouvement le long d'une ligne droite. C'est un système à un degré de liberté, dont le mouvement est décrit par une équation du second ordre. On a besoin d'une coordonnée généralisée de position pour caractériser ce mouvement, s , le déplacement le long de la ligne. La vitesse généralisée est la dérivée première de cette coordonnée par rapport au temps. En insérant cette coordonnée généralisée dans l'équation de Lagrange (41a), on peut écrire :

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{s}} \right) - \frac{\partial L}{\partial s} = 0 \quad (26)$$

Après spécification des conditions initiales, l'intégration donne une fonction unique décrivant une courbe dans l'espace des états.

Particule décrivant un mouvement harmonique simple Prenons maintenant l'exemple d'une particule unique décrivant un mouvement harmonique simple. Ici, $L = T - V = \frac{1}{2}m\dot{s}^2 - \frac{1}{2}ks^2$. Donc, $\frac{\partial L}{\partial s} = -ks$, $\frac{\partial L}{\partial \dot{s}} = m\dot{s}$ et $\frac{d}{dt}(\frac{\partial L}{\partial \dot{s}}) = m\ddot{s}$. On obtient la même équation du mouvement que celle obtenue dans la formulation newtonienne, avec la coordonnée newtonienne x et sa dérivée par rapport au temps remplacées par la coordonnée généralisée de position s et sa dérivée par rapport au temps :

$$m\ddot{s} = -ks \quad (27)$$

Notons que cette équation a la même forme, que s soit exprimé en coordonnées rectangulaires ou angulaires. (27) est l'équation du mouvement d'une particule unique libre sous l'action d'un mouvement harmonique simple, c'est-à-dire sous l'action d'une force de rappel qui est une fonction de la particule, quelles que soient les coordonnées de ce mouvement.

⁵⁴En considérant le mouvement d'une particule au moyen de coordonnées polaires planes, on retrouve également l'équation du moment de la force, qui est une généralisation de l'équation newtonienne. Voir (Goldstein, 1950/2002, pp. 26-27).

Pendule simple J'ai présenté, à la section 1.3.1, la manière dont la formulation newtonienne permet de résoudre le problème du pendule simple. On avait vu que l'équation de Newton appliquée à ce problème n'avait pas la forme simple de la loi de Hooke, car l'expression de la force dépendait du sinus de l'angle d'oscillation ; dans de tels cas, l'utilisation de coordonnées polaires peut s'avérer plus simple. J'avais alors signalé le fait que la forme de la loi de Newton, en raison de son caractère vectoriel, change avec un tel changement de coordonnées.

En mécanique lagrangienne, la forme des équations est indépendante d'un changement de coordonnées. Le seul trait pertinent, pour la représentation lagrangienne du mouvement du pendule, est l'angle θ . En prenant $\dot{\theta}$ comme vitesse généralisée⁵⁵, on obtient la même équation que dans le cadre newtonien, avec la coordonnée généralisée θ et sa dérivée à la place de x et de sa dérivée :

$$-g \sin \theta = l \ddot{\theta} \quad (28)$$

Comme c'était le cas dans la formulation newtonienne, l'équation du mouvement du pendule (28) n'a pas la même forme que pour le mouvement harmonique simple (27) : la première contient le sinus de la coordonnée de position. Cependant, les équations de Lagrange (41a) ne changent pas de forme en coordonnées polaires, alors que les équations de Newton, elles, changent. L'équation qui décrit le mouvement du pendule est bien la même dans le cadre newtonien et dans le cadre lagrangien. Mais, en raison du fait qu'elle met en jeu des quantités scalaires et non vectorielles, la loi dynamique de laquelle on dérive cette équation ne change pas de forme, en mécanique lagrangienne, avec un changement de coordonnées.⁵⁶

En conséquence, pour résoudre tous ces problèmes, il n'est pas nécessaire d'écrire les équations des forces, et, par conséquent, il n'est pas nécessaire de connaître les forces de contraintes. On aurait pu commencer par utiliser des coordonnées rectangulaires, calculer le lagrangien dans ces termes, et ensuite utiliser les équations de transformation entre les systèmes de coordonnées pour obtenir une solution. Mais cela est inutile. Dans le cadre lagrangien, le fait que la trajectoire d'une particule soit rectiligne ou courbe n'a pas d'importance. Le trait pertinent, dans les cas présentés ci-dessus, est le caractère unidimensionnel du mouvement, ce qui indique que l'on peut le représenter au moyen d'une unique coordonnée généralisée. On peut ensuite choisir la coordonnée qui est « naturelle » au système (rectiligne, polaire, etc.), et de l'insérer dans l'équation de Lagrange (41a) : dans cette équation, on remplace simplement q et \dot{q} par s et \dot{s} , ou θ et $\dot{\theta}$.

⁵⁵ $L = T - V = \frac{1}{2}m(l\dot{\theta})^2 + mgl \cos \theta$. Donc, $\frac{\partial L}{\partial \theta} = -mgl \sin \theta$, $\frac{\partial L}{\partial \dot{\theta}} = ml^2 \dot{\theta}$, et $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) = ml^2 \ddot{\theta}$.

⁵⁶ Sur ce point, voir (North, 2009a).

Les pendules liés Cette particularité de la mécanique analytique la rend beaucoup plus facile d'application dans les cas de systèmes contraints pour lesquels la formulation newtonienne est pratiquement inutilisable (voir page 50). Par exemple, dans le cas des deux pendules liés (voir figure 4, page 50), la formulation newtonienne est très peu pratique : elle nécessite de faire intervenir la tension des fils, dont on ne se débarrasse ensuite qu'au prix de manipulations fastidieuses.

En revanche, ce système est très simple à décrire au moyen de la formulation lagrangienne : il a deux degrés de liberté. Les coordonnées généralisées sont donc les angles θ_1 et θ_2 . On écrit sans difficulté les valeurs des positions et des vitesses des deux masses en fonction des coordonnées généralisées et des longueurs l_1 et l_2 des deux pendules. On en déduit l'expression des énergies cinétique et potentielle du système⁵⁷ :

$$T = T_1 + T_2 \quad (29)$$

$$T_1 = \frac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 \quad (30)$$

$$T_2 = \frac{1}{2}m_2 [l_1^2 \dot{\theta}_1^2 + l_2^2 \dot{\theta}_2^2 + 2l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2] \quad (31)$$

$$U = U_1 + U_2 \quad (32)$$

$$U_1 = -m_1 g l_1 \cos(\theta_1) \quad (33)$$

$$U_2 = -m_2 g l_2 \cos(\theta_2) - m_2 g l_1 \cos(\theta_1) \quad (34)$$

On peut alors écrire l'expression du lagrangien L et obtenir le système d'équations lagrangiennes décrivant le mouvement du pendule. En raison de la présence de termes en $\cos\theta$, ces équations ne sont pas linéaires. On doit donc procéder à une approximation linéaire, valide dans le domaine des petites oscillations. Après linéarisation, les équations pour θ_1 et θ_2 s'écrivent :

$$(m_1 + m_2)l_1 \ddot{\theta}_1 + m_2 l_2 \ddot{\theta}_2 + (m_1 + m_2)g\theta_1 = 0 \quad (35)$$

$$l_2 \ddot{\theta}_2 + l_1 \ddot{\theta}_1 + g\theta_2 = 0 \quad (36)$$

On résout ce système de deux équations linéaires à coefficients constants par les méthodes standard. Les solutions sont des combinaisons linéaires de solutions en $\exp(-i\omega t)$. On obtient une équation bicarrée en ω . Les solutions s'écrivent $\pm\omega_1$ et $\pm\omega_2$. La solution générale est une superposition de deux mouvements oscillatoires aux fréquences ω_1 et ω_2 .

À aucun moment, il n'a été nécessaire de prendre en compte les forces en présence ; la seule connaissance des angles θ_1 et θ_2 permet de résoudre le problème sans difficulté.

⁵⁷La déduction qui suit est tirée du cours de Jean-Michel Raimond (2000).

2.4.2 Limites pratiques de la formulation lagrangienne

On vient de voir que la formulation lagrangienne présente de nombreux avantages par rapport à la formulation newtonienne, nous offrant un procédé très puissant pour établir les équations du mouvement d'un système contraint pour lequel on peut définir un lagrangien. Notons cependant que cette dernière condition implique une restriction des problèmes que la formulation lagrangienne permet de résoudre. En effet, pour les systèmes mettant en jeu des forces qui ne dérivent pas d'un potentiel, rendant impossible la définition d'un lagrangien, la formulation newtonienne reste la seule utilisable. C'est le cas, par exemple, des systèmes mettant en jeu des forces de frottement.

En outre, on se heurte souvent, lorsque l'on résout des problèmes au moyen de la formulation lagrangienne, à la difficulté suivante : les équations de Lagrange, étant des différentielles du second ordre – tout comme les équations newtoniennes – ne sont pas toujours intégrables.⁵⁸ La formulation hamiltonienne, dont je vais à présent exposer les aspects élémentaires, permet de surmonter ce problème.

3 Mécanique analytique – Formulation hamiltonienne

La formulation hamiltonienne de la mécanique fait partie, avec la formulation lagrangienne, de ce que l'on appelle la « mécanique analytique ». Comme elle, elle repose sur la représentation de la dynamique des systèmes physiques à l'aide de coordonnées généralisées ; les équations hamiltoniennes du mouvement, comme les équations lagrangiennes, sont déductibles d'un principe variationnel.

Pas plus que la formulation lagrangienne ne « remplace » la formulation newtonienne, qui reste beaucoup plus facile d'utilisation dans de nombreux cas, la formulation hamiltonienne ne remplace la formulation lagrangienne. On l'introduit cependant la plupart du temps après avoir présenté la formulation lagrangienne, par rapport à laquelle elle présente plusieurs avantages. Après avoir présenté quelques aspects élémentaires de la formulation hamiltonienne, je reviendrai, dans la section 4 du présent chapitre, sur la comparaison des trois formulations présentées.

⁵⁸Cependant, une intégration partielle des équations lagrangiennes permet d'obtenir des équations différentielles du premier ordre appelées « intégrales premières du mouvement », qui fournissent des renseignements sur la dynamique du système étudié, même si sa trajectoire n'est pas totalement connue. Elles permettent notamment d'établir un lien entre les propriétés du système et les lois de conservation (voir l'annexe A.11).

3.1 Transformations de Legendre et équations hamiltoniennes du mouvement

Un des aspects essentiellement nouveaux de la formulation hamiltonienne de la mécanique réside dans le fait que ses équations du mouvement sont, contrairement à celles de Newton et de Lagrange, des équations différentielles *du premier ordre*. Cela a pour conséquence, entre autres, de faciliter la résolution des problèmes, car les équations du premier ordre s'intègrent aisément et de façon numériquement stable.

Aux n équations différentielles du second ordre nécessaires, dans le cadre lagrangien, pour décrire le mouvement d'un système à n degrés de liberté, la formulation hamiltonienne permet de substituer $2n$ équations du premier ordre, obtenues par transformation des n équations lagrangiennes.⁵⁹ Cette nouvelle manière de décrire la dynamique des systèmes physiques repose sur l'utilisation de coordonnées généralisées d'un type différent de celles utilisées dans la formulation lagrangienne. En effet, les variables indépendantes ne sont plus les positions et vitesses généralisées, mais les positions et quantités de mouvement (moments) généralisées. On définit les quantités de mouvement généralisées p_i comme suit :

$$p_i = \frac{\partial L(q_i, \dot{q}_i, t)}{\partial \dot{q}_i} \quad (37)$$

Les équations de Lagrange peuvent alors être écrites sous la forme

$$\dot{p}_i = \frac{\partial L}{\partial q_i} \quad (38)$$

ce qui fait apparaître L comme une fonction des q_i et des \dot{q}_i , dont les dérivées partielles respectives sont les \dot{p}_i et les p_i . La différentielle totale du lagrangien s'écrit alors :

$$dL = \sum_i p_i d\dot{q}_i + \sum_i \dot{p}_i dq_i \quad (39)$$

La technique des transformations de Legendre⁶⁰ permet de passer de L à une fonction H dont la différentielle s'exprime en fonction des q_i et des p_i . On appelle « fonction de Hamilton » ou « hamiltonien » cette fonction, définie ainsi :

$$H(p, q, t) = \sum_i \dot{q}_i p_i - L(q, \dot{q}, t) \quad (40)$$

⁵⁹ Je ne considérerai ici que les systèmes mécaniques holonomes et admettant un potentiel généralisé.

⁶⁰ Cette technique inventée par le mathématicien français Legendre (1752-1833) permet de changer une fonction donnée d'un ensemble de variables données en une nouvelle fonction d'un nouvel ensemble de variables, les anciennes et nouvelles variables étant reliées entre elles par des transformations de point. Ici, le procédé permet de rendre le lagrangien linéaire dans les vitesses, en doublant le nombre des variables mécaniques. Voir Goldstein, 1950/2002, pp. 334-338) et (Lanczos, 1970, pp. 161-164).

La différentielle de l'hamiltonien est

$$dH = \sum_i \dot{q}_i dp_i - \sum_i \dot{p}_i dq_i \quad (41)$$

H est donc une fonction des p_i et q_i dont les dérivées partielles respectives sont les \dot{q}_i et les \dot{p}_i . Les n équations de Lagrange peuvent alors être remplacées par $2n$ équations différentielles du premier ordre, qui leur sont équivalentes, appelées les *équations canoniques de Hamilton* :

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \quad (42)$$

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i \quad (43)$$

Les positions et les quantités de mouvement généralisées sont appelées « variables conjuguées ».

Comme dans la formulation lagrangienne, une fonction scalaire unique permet de prédire le comportement du système : c'est l'hamiltonien H , typiquement égal à l'énergie totale du système.⁶¹ Pour un état initial donné, il permet de déterminer une histoire unique du système, parmi ses trajectoires possibles.⁶²

3.2 Dédution des équations de Hamilton d'un principe variationnel

Les équations hamiltoniennes, obtenues par transformations des équations lagrangiennes, peuvent également, comme ces dernières, être obtenues à partir d'un principe

⁶¹Pour des systèmes conservatifs aux contraintes holonomes, l'hamiltonien coïncide avec l'énergie totale. En revanche, pour des systèmes où le lagrangien dépend explicitement du temps, l'hamiltonien ne correspond plus à l'énergie mécanique totale mais permet toujours d'écrire les équations du mouvement.

⁶²Le formalisme hamiltonien se prête particulièrement bien au traitement de problèmes mettant en jeu des coordonnées cycliques. On dit qu'une coordonnée q_i est cyclique ou ignorable quand le lagrangien ou l'hamiltonien ne la contient pas (bien qu'il puisse contenir la vitesse \dot{q}_i ou la quantité de mouvement p_i correspondante). Voir (Goldstein, 1950/2002, pp. 343-347) et l'annexe A.11. On peut notamment transposer dans la formulation hamiltonienne les théorèmes de conservation de la quantité de mouvement par une simple substitution de H à L . En effet, si q_i est une coordonnée cyclique, le lagrangien et donc l'hamiltonien ne dépendent pas explicitement de q_i . On peut alors déduire immédiatement des équations de Hamilton que p_i est une constante du mouvement. On retrouve dans le cadre de la formulation hamiltonienne ce qu'on constate dans celui de la formulation lagrangienne : si L n'est pas une fonction explicite de t , alors H est une constante du mouvement (voir l'annexe A.11).

variationnel. Il suffit pour cela d'exprimer l'action A (définie par $A = \int_{t_1}^{t_2} L dt$) en fonction de l'hamiltonien. En vertu de la définition (40) de l'hamiltonien, on peut ainsi écrire :

$$A = \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H \right) dt \quad (44)$$

Le principe de Hamilton devient

$$\delta A = \delta \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H \right) dt = 0 \quad (45)$$

On appelle parfois l'équation (45) le *principe de Hamilton modifié*. Il nous dit que l'accroissement de l'action dans un accroissement infinitésimal doit s'annuler. En considérant la trajectoire effectivement suivie, définie par $q_i(t)$ et $p_i(t)$, et une trajectoire infiniment proche, définie à chaque instant par $q_i(t) + \delta q_i(t)$ et $p_i(t) + \delta p_i(t)$, on impose aux deux trajectoires de coïncider à l'instant initial et à l'instant final : $\delta q_i(1) = \delta q_i(2) = 0$. L'accroissement de l'action peut s'écrire :

$$\delta A = \int_{t_1}^{t_2} \left(\sum_i (\delta p_i \dot{q}_i + p_i \delta \dot{q}_i) - \delta H \right) dt \quad (46)$$

L'accroissement de H entre la trajectoire de référence et la trajectoire variée s'écrit simplement :

$$\delta H = \sum_i \frac{\partial H}{\partial q_i} \delta q_i + \frac{\partial H}{\partial p_i} \delta p_i \quad (47)$$

En intégrant par parties, on peut finalement en déduire les équations de Hamilton.

3.3 Transformations canoniques

Pour finir, signalons que la formulation hamiltonienne permet d'opérer un changement de coordonnées très général, de la forme

$$\begin{aligned} q_i &\rightarrow Q_i(q_i, p_i, t) \\ p_i &\rightarrow P_i(q_i, p_i, t) \end{aligned} \quad (48)$$

Les nouvelles positions sont définies au moyen des anciennes positions, mais aussi au moyen des anciens moments. Un tel changement de variables laisse toute liberté⁶³ dans le choix des coordonnées, afin d'obtenir des équations intégrables. Cela permet dans de traiter des problèmes qui sont parfois insolubles dans le cadre lagrangien.

⁶³Il faut cependant exiger qu'il soit réversible, afin que l'état du système soit déterminé de façon univoque par les nouvelles coordonnées. Cela implique que les Q_i et les P_i soient elles aussi indépendantes.

S'il existe un nouvel hamiltonien $H'(Q_i, P_i, t)$ qui donne les équations hamiltoniennes pour les nouvelles variables, on appelle cette transformation « canonique⁶⁴ » et on peut l'utiliser pour résoudre des problèmes de mécanique. Pour que le principe de Hamilton, de la forme de l'équation (45), formulé à l'aide des anciennes variables et à l'aide des nouvelles variables donne les mêmes équations du mouvement, il suffit que les deux quantités intégrées ne diffèrent que de la dérivée totale par rapport au temps d'une fonction F , appelée « fonction génératrice ». S'il existe une telle fonction F , la transformation est canonique. Une fois la fonction F donnée, les équations de transformation (48) sont complètement définies. On dispose ainsi d'une méthode très puissante de résolution des problèmes de mécanique.

4 En quel sens peut-on dire que les formulations de la mécanique sont équivalentes ?

Quelles conclusions peut-on tirer de ce qui précède pour l'analyse de la notion de théorie ? Dans cette section, je reviens d'abord sur les éléments qui permettent d'affirmer que ces trois formulations sont équivalentes, et je dégage les présupposés qui sous-tendent l'affirmation courante selon laquelle elles sont l'expression d'une seule et même théorie (section 4.1). Ensuite, je dresse un bilan des principales différences entre ces formulations, mises en évidence par l'exposé qui précède (section 4.2). Enfin, je me concentre sur un type particulier de différences, celles qui concernent la représentation mathématique offerte par chacune des formulations ; j'insiste en particulier sur le lien entre ces différences « objectives » (dont on considère cependant, généralement, qu'elles n'affectent pas l'équivalence des formulations) et l'utilisation de chacune de ces formulations, en pratique, pour résoudre des problèmes (section 4.3). L'analyse de ces différences mathématiques et de leur contrepartie pratique me permet d'avancer un premier argument visant à montrer que les présupposés examinés à la section 4.1 doivent être abandonnés si l'on veut rendre compte de la double fonction représentative et inférentielle des théories, et d'ouvrir la voie d'une approche des théories qui adopte la perspective de leurs utilisateurs (section 4.4).

4.1 Équivalence *en principe*

Les trois formulations de la mécanique présentées ci-dessus sont généralement considérées comme équivalentes. Une transformation mathématique permet de passer de n'importe laquelle des trois sortes d'équations à chacune des deux autres. Les

⁶⁴Ce terme est dû à Jacobi.

équations de Lagrange peuvent être tirées du principe de d'Alembert (page 56), qui peut lui-même être considéré comme une reformulation des lois de Newton (page 55) ; elles sont, de plus, équivalentes au principe de Hamilton (page 60), déductible lui aussi du principe de d'Alembert (page 59). Les équations de Hamilton sont elles-mêmes des transformations mathématiques des équations de Lagrange (page 67), et on peut les déduire du principe de Hamilton modifié par la même transformation mathématique (page 68). Enfin, on retrouve la forme des équations newtoniennes à partir des équations lagrangiennes et hamiltoniennes (page 62).

Certes, ces différentes déductions impliquent souvent des restrictions ; par exemple, la déduction du principe de Hamilton à partir du principe de d'Alembert suppose que l'on se restreigne à des systèmes holonomes et à des forces dérivant d'un potentiel. La mécanique newtonienne reste la seule utilisable pour l'étude des forces de frottement. Ces restrictions n'entraînent cependant pas de contradiction entre ces différents principes. Pour bien comprendre ce que cela signifie, reprenons l'exemple de la mécanique newtonienne et de la mécanique einsteinienne. Cette dernière est une généralisation de la mécanique newtonienne, au sens où la mécanique newtonienne devient un cas particulier de la mécanique einsteinienne, mais, en outre, elle la contredit : elle montre que la mécanique newtonienne n'est qu'approximativement vraie pour un certain domaine de phénomènes (les mouvements où les vitesses sont très inférieures à celle de la lumière). En revanche, les différents principes de la mécanique analytique n'impliquent jamais de conséquences contradictoires. Pour les domaines où ils sont applicables, leurs prédictions empiriques ne peuvent pas entrer en contradiction.

C'est la raison pour laquelle on les considère généralement comme l'expression d'une seule et même théorie. La reformulation analytique est habituellement considérée comme une généralisation, qui permet d'étendre le domaine d'application de la mécanique newtonienne et la rend, ce faisant, plus « puissante ». À strictement parler, elle en est un développement mathématique. L'inter-déductibilité entre les principes de la mécanique, en garantissant *a priori* qu'il ne peut pas y avoir de contradiction entre les conséquences empiriques des différentes équations, permet donc d'affirmer qu'elles sont, en un sens, équivalentes. Appelons cette équivalence une équivalence *logique*⁶⁵ ou *en principe*. Si l'on adopte cette notion d'équivalence pour évaluer les

⁶⁵Comme je l'ai signalé à la note 1 du présent chapitre, parler d'équivalence logique implique que l'on précise le langage dans lequel on évalue cette équivalence. Chercher à donner un critère précis d'équivalence logique pour les différentes formulations de la mécanique m'éloignerait de mon propos, et n'apporterait rien à mon argument, dans la mesure où ce n'est pas sur le front de la caractérisation formelle des relations inter-théoriques (ou inter-formulations) que je souhaite avancer mes arguments. Au contraire, mon but est de montrer que, même en supposant que les formulations

relations entre différentes formulations, alors le contenu d'une théorie est entièrement défini par l'ensemble de ses conséquences déductives *en principe*, indépendamment de la manière dont ces conséquences sont atteintes par les agents, *en pratique*.

Mon but est de montrer que cette notion d'équivalence est mal choisie quand il s'agit d'analyser les théories scientifiques. Je veux dire par là que, quels que soient les critères formels que l'on peut établir pour montrer cette équivalence, affirmer sur cette base que le contenu des formulations de la mécanique est strictement identique implique une conception des théories qui ne rend pas compte de leur double fonction représentative et inférentielle. Revenons à présent sur les différences que mon exposé des formulations de la mécanique a mises en évidence.

4.2 Bilan des différences

En premier lieu, chacune des trois formulations de la mécanique présentées ci-dessus offre une représentation des phénomènes du mouvement au moyen de différents *principes* généraux (deuxième loi de Newton, principes de d'Alembert et de Hamilton, principe de Hamilton modifié) desquels on peut tirer des équations différentielles permettant de prédire et d'expliquer le comportement de systèmes mécaniques. Le principe fondamental de la dynamique newtonienne est lui-même une *équation différentielle*, dont le concept central est celui de force, exprimé au moyen d'une quantité vectorielle. Il exprime la relation qu'entretiennent à chaque instant les forces appliquées au système et l'accélération de ce système; en connaissant ces forces et les conditions initiales du système, on peut écrire les équations du mouvement, qui ont la forme du principe fondamental de la dynamique. Le principe de Hamilton est un *principe variationnel* qui exprime une condition (celle de la minimisation d'une intégrale) sur le mouvement des systèmes pris dans sa globalité, dont on peut déduire des équations différentielles permettant de prédire et d'expliquer le comportement des systèmes mécaniques. Le concept central de ce principe et de ces équations est celui d'énergie, exprimé au moyen d'une quantité scalaire.

Si l'on considère ces principes comme les hypothèses que la théorie exprime à propos des phénomènes, et au moyen desquels elle nous permet de les comprendre, il semble, en un sens qui reste à analyser, que ces trois formulations *ne disent pas la même chose du monde*. L'une parle de forces et offre une représentation locale du

de la mécanique puissent être dites équivalentes en un sens strict, des différences qui ne vaudraient pas comme des différences « logiques » sont importantes pour l'analyse de l'activité théorique. On trouvera cependant, au chapitre 5, une critique des approches formelles en philosophie des sciences sur leur propre terrain. Je me contente, ici, d'appeler « équivalence logique » une équivalence empirique prouvée *a priori*, c'est-à-dire une équivalence telle qu'on ne peut pas déduire de conséquences contradictoires des différentes formulations, quel que soit l'état du monde.

mouvement ; les deux autres parlent d'énergie, et en offrent une représentation globale.⁶⁶ En un sens pré-analytique et banal de la notion de signification, l'affirmation selon laquelle la force et l'accélération sont proportionnelles n'a pas la même signification que l'affirmation selon laquelle l'intégrale de la différence de l'énergie cinétique et de l'énergie potentielle est stationnaire. De même, concevoir les phénomènes du mouvement en termes d'énergie ou en termes de force sont deux manières différentes de se représenter le monde.

Certes, une fois définis les concepts d'énergie et de force, chacune de ces deux affirmations est déductible de l'autre, et l'on peut raisonnablement affirmer que la compréhension de la mécanique classique – et de ce qu'elle dit des phénomènes – implique la compréhension de cette inter-déductibilité. Mais cela ne supprime pas *ipso facto* la différence conceptuelle entre le principe fondamental de la dynamique et le principe de Hamilton. Un des buts de mon travail, dans son ensemble, est d'approfondir l'analyse de ce que l'on entend ici par « différence conceptuelle », et de montrer que toute entreprise visant à la réduire à une différence purement « psychologique » qu'une analyse philosophique de la notion de théorie ne doit pas prendre en compte est condamnée à manquer des aspects essentiels de cette notion. Pour l'heure, je me contente de souligner cette différence, dont j'approfondirai l'analyse dès le chapitre 3, et je me concentre, dans ce chapitre, sur un autre type de différences entre les formulations de la mécanique.

Outre cette différence des principes et des concepts, les formulations de la mécanique se distinguent en effet les unes des autres par la *forme des équations* – déduites de ces principes – au moyen desquelles elles permettent de résoudre des problèmes de mécanique. Ces équations gouvernent la dynamique des systèmes mécaniques en représentant leur configuration dans un espace géométrique, au moyen d'un certain type de coordonnées (coordonnées cartésiennes ou polaires, coordonnées généralisées).

On a vu que les changements dans la représentation mathématique de l'état d'un système et dans la forme des équations qui lui sont associées ont d'importantes conséquences pratiques pour la résolution de problèmes mécaniques. Dans certains cas, l'utilisation d'une certaine formulation rend impossible la *mise en équations* du problème ; dans d'autres cas, c'est la *résolution de ces équations* qui est impossible. Ainsi, un changement dans la *représentation mathématique* d'un problème entraîne

⁶⁶Signalons d'emblée une difficulté supplémentaire, que l'on retrouvera au cours de l'analyse du chapitre 3 : si l'on prête attention au langage mathématique utilisé dans ces différents principes, la frontière entre formulations newtonienne et variationnelle n'est plus la même : en effet, le principe de d'Alembert est un principe différentiel, comme le PFD de Newton, alors que le principe de Hamilton est variationnel.

dans certains cas un changement dans l'*accessibilité*⁶⁷ des solutions de ce problème, dans le type d'information dont il faut disposer à propos des systèmes étudiés (par exemple, connaître les différentes forces appliquées au système), et dans les processus inférentiels que la recherche de ces solutions requiert. Dans la prochaine section, je me propose d'approfondir l'examen de la relation entre ces aspects pratiques et le type de représentation mathématique mis en œuvre par chacune des formulations.

4.3 Représentation mathématique et différences pratiques

On appelle « espace des états » l'espace mathématique dans lequel chacun des points représente un des *états possibles* du système entier. L'espace des états ne coïncide pas avec l'espace physique ; c'est un espace beaucoup plus abstrait. Dans le cas d'une particule unique en mouvement libre dans un espace à trois dimensions, chacun de ses états possibles est déterminé par six coordonnées (trois pour les positions et trois pour les vitesses) ; son espace des états a donc six dimensions. Pour un système à n particules libres de contraintes, chaque état possible étant déterminé par $6n$ coordonnées, l'espace des états de ce système a $6n$ dimensions. Différentes courbes ou trajectoires dans cet espace représentent des *histoires* possibles du système.

Les lois de la dynamique, exprimées par les équations newtoniennes, lagrangiennes ou hamiltoniennes, disent quelles courbes dans l'espace des états sont des histoires possibles. Elles imposent des contraintes *dynamiques* sur la trajectoire des systèmes, dont l'état est représenté dans un certain espace. La description du mouvement d'un système au moyen d'une équation suppose que son espace des états ait une certaine structure géométrique. Or, les différentes formulations de la mécanique n'assignent pas la même structure géométrique à l'espace dans lequel les équations dynamiques décrivent l'évolution de l'état des systèmes. Comme je vais rapidement le montrer, chacune des formulations assigne à l'espace des états une structure plus générale que la précédente ; c'est la raison pour laquelle on peut parler de généralisations successives entre formulations. Ces généralisations ont une contrepartie pratique : elles s'accompagnent d'une plus grande liberté dans le choix des variables utilisées pour représenter le système et écrire les équations qui gouvernent sa dynamique.⁶⁸

En *mécanique newtonienne*, l'état d'un système à chaque instant est caractérisé par deux ensembles de coordonnées indépendantes, celles des positions et des quantités de mouvement (produit de la masse et de la vitesse) de chacune de ses particules. Pour un mouvement libre dans un espace à trois dimensions, l'état d'un système

⁶⁷C'est un des thèmes centraux de l'ouvrage de Paul Humphreys (2004) ; on le retrouvera à plusieurs reprises dans le présent travail.

⁶⁸L'analyse qui suit est partiellement empruntée à Jill North (2009a,b).

composé de n particules est caractérisé par un total de $6n$ coordonnées.

L'évolution du système dans son espace des états est gouvernée par le principe fondamental de la dynamique, $\mathbf{F} = m \frac{d\mathbf{P}}{dt}$.⁶⁹ Il s'agit d'une équation *vectorielle*, qui est invariante par transformation galiléenne. La deuxième loi de Newton continue à valoir dans d'autres systèmes de coordonnées, mais sa forme change dans une transformation, par exemple, de coordonnées cartésiennes en coordonnées polaires. Les vecteurs sont des objets indépendants des coordonnées, mais leurs composantes changent avec le système de coordonnées.

La *contrepartie pratique* de l'utilisation de ce type de représentation est la nécessité de spécifier les composantes de chacune des forces appliquées au système dans les coordonnées qu'on aura choisies afin d'écrire les équations vectorielles gouvernant la dynamique du système : les vecteurs sont des objets indépendants des coordonnées, mais leurs composantes changent avec le système de coordonnées ; pour résoudre un problème avec la loi de Newton, il faut connaître les composantes des forces dans le système choisi. Les équations newtoniennes du mouvement, en raison même de leur nature vectorielle, privilégient un certain type de coordonnées (North, 2009a, p. 22) ; utiliser la mécanique newtonienne implique d'admettre que le mouvement de tout système peut être décrit dans un espace vectoriel à l'aide de coordonnées cartésiennes. *La structure géométrique de l'espace de représentation de la mécanique newtonienne est vectorielle.*

En *mécanique lagrangienne*, l'état d'un système est aussi spécifié par deux ensembles de coordonnées : les positions (ou coordonnées⁷⁰) généralisées et leurs dérivées par rapport au temps, les vitesses généralisées. Pour un système à n degrés de liberté, on aura $6n$ coordonnées indépendantes (les n coordonnées généralisées qui représentent le système dans l'espace de configuration⁷¹ et leurs dérivées par rapport au temps). Chaque état possible du système est donc représenté par un point dans un espace à $6n$ dimensions, qui définit une paire position généralisée-vitesse généralisée pour l'ensemble du système. Cet espace comprend donc un ensemble de configurations possibles et un espace dans lequel sont représentées leurs dérivées par rapport au temps. Il s'agit du *faisceau tangent* de l'espace de configuration : l'espace

⁶⁹On peut donner une « super-version » de l'équation newtonienne du mouvement, regroupant toutes les équations pour les différentes particules dans un espace tri-dimensionnel, qui gouverne l'évolution du point représentant l'état du système entier dans l'espace à $6n$ dimensions.

⁷⁰Voir note 30.

⁷¹Il faut veiller à ne pas confondre l'espace de configuration et l'espace des états : l'espace de configuration, en mécanique lagrangienne, a autant de dimensions que le système a de degrés de liberté ; il permet de représenter la configuration du système à l'aide des positions généralisées. Autrement dit, il correspond à l'espace euclidien à trois dimensions dans lequel la position d'un système est représentée dans le cadre newtonien.

de configuration à $3n$ dimensions joint à l'espace tangent à $3n$ dimensions à chaque point (la figure 6 représente le faisceau tangent d'un cercle unidimensionnel).

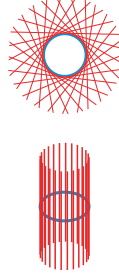


FIG. 6 – Le faisceau tangent d'une variété (espace topologique abstrait) – en l'occurrence, un cercle – est l'ensemble de ses espaces tangents (représentés dans la figure du haut) assemblés de manière homogène et sans chevauchement (figure du bas).

L'évolution d'un système dans son espace des états est ici gouvernée par une équation *scalaire* et non vectorielle. La forme de cette équation n'est pas modifiée par un changement de coordonnées, car le lagrangien L , qui est un scalaire, de la dimension d'une énergie, est indépendant du système de coordonnées choisi. En conséquence, l'espace des états d'un système reste le même, quelles que soient les coordonnées choisies.⁷²

La *contrepartie pratique* de l'utilisation de ce type de représentation mathématique est la liberté dans le choix des variables utilisées pour définir le lagrangien, ce qui permet de décrire simplement l'état d'un système contraint, pour lequel la description à l'aide d'équations vectorielles représentant les forces est pratiquement impossible. Étudier la structure géométrique de l'espace des états de la formulation lagrangienne nécessiterait d'avoir recours aux outils de la géométrie moderne, et cela dépasserait les limites de ce travail (et mes compétences). Signalons seulement que la mécanique lagrangienne assigne à l'espace des états une structure *métrique*. La structure de la mécanique lagrangienne est un espace riemanien, muni d'une métrique

⁷²Prenons l'exemple simple d'une particule en mouvement le long d'une ligne droite. Son espace des états est un espace à deux dimensions : le faisceau tangent d'une ligne ou d'un cercle unidimensionnel. Le cercle de la figure 6 représente les positions généralisées q . Les lignes (fibres) représentent les vitesses généralisées. Chaque point représente un état possible. Différentes courbes dans cet espace représentent des histoires possibles. Il y a plusieurs trajectoires possibles passant par un point pour une valeur q donnée. Une fois la valeur \dot{q} spécifiée elle aussi, une trajectoire unique est déterminée. On voit que l'espace des états pour une particule en mouvement unidimensionnel sera le même, que q soit une coordonnée linéaire ou angulaire. Revenons à l'exemple du pendule simple : la seule coordonnée que l'on ait besoin de spécifier est θ , avec $\dot{\theta}$ comme vitesse généralisée. L'espace des états est de nouveau celui de la figure 6, avec le cercle représentant θ et les fibres représentant $\dot{\theta}$.

riemanienne⁷³ ; utiliser la mécanique lagrangienne, c'est admettre que le mouvement de tout système peut être décrit dans un tel espace.

Enfin, en *mécanique hamiltonienne*, le nombre de variables par degré de liberté d'un système double, mais ces variables sont indépendantes. Rappelons en effet que les deux ensembles de variables nécessaires à décrire l'état d'un système dans la formulation lagrangienne ne sont pas indépendantes : les vitesses généralisées sont les dérivées des positions généralisées par rapport au temps. L'espace des états de la mécanique hamiltonienne est appelé « *espace des phases* ». C'est le *faisceau cotangent*⁷⁴ de l'espace de configuration (voir figure 7). C'est, ici aussi, un espace à $6n$ dimensions. Chaque point y représente une paire position généralisée-moment généralisé pour tout le système. Ici encore, la dynamique du système est gouvernée par une fonction scalaire. L'espace des états – ou espace des phases – de la mécanique hamiltonienne, contrairement à l'espace des états lagrangien, n'a pas de structure métrique : il a une structure *symplectique*.⁷⁵

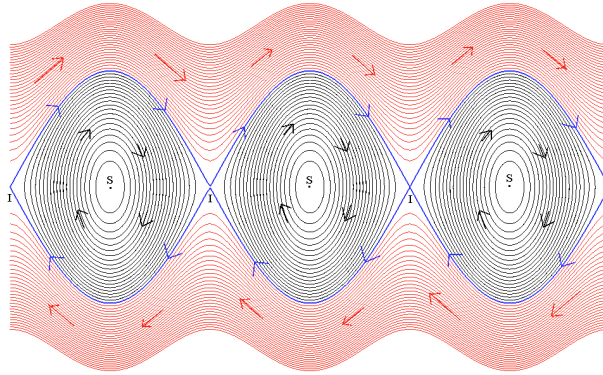


FIG. 7 – L'espace des phases du pendule simple est le faisceau cotangent de l'espace de configuration.

Les différences des structures géométriques assignées à l'espace des états des systèmes par chacune des formulations n'impliquent et ne peuvent impliquer aucune différence empirique : une fois les contraintes dynamiques imposées par les équations

⁷³Cela apparaît quand on considère les quantités invariantes des équations lagrangiennes sous un certain ensemble de transformations (North, 2009b). La transformation d'un ensemble d'équations lagrangiennes en un autre préserve une structure locale, sous la forme d'une forme différentielle quadratique des \dot{q}_i ; il s'agit du carré d'un élément de courbe d'un espace riemanien (voir Butterfield, 2004, 2006 et Lanczos, 1970, pp. 17-24).

⁷⁴Le faisceau cotangent d'un espace topologique est l'ensemble de ses espaces cotangents.

⁷⁵Une structure symplectique détermine un élément de volume, et non une mesure de distance. La quantité invariante sous transformation canonique des équations hamiltoniennes est une forme symplectique (voir North, 2009b, p. 18 ; Butterfield, 2006 ; Arnold, 1978).

tions, les trajectoires possibles sont équivalentes dans les différentes formulations. Cependant, la structure des espaces sous-jacents aux différentes formulations impose des contraintes différentes sur les mouvements *géométriquement* possibles, c'est-à-dire possibles *avant que les contraintes dynamiques soient imposées*.

Ces considérations montrent que, quand bien même les formulations de la mécanique sont logiquement équivalentes au sens, mentionné précédemment, où leur équivalence empirique est démontrable *a priori* (indépendamment de l'expérience), il y a un sens où elles ne sont pas *mathématiquement équivalentes* : elles n'assignent pas la même structure géométrique aux espaces dans lesquelles elles décrivent le mouvement des systèmes. Chacun de ces espaces a une structure plus générale ou abstraite : pour reprendre les termes de Jill North (2009b), l'espace hamiltonien a *moins* de structure que l'espace lagrangien, lequel en a moins, à son tour, que l'espace newtonien.⁷⁶ En effet, une structure symplectique définit un élément de volume, mais pas de distance ; une structure métrique définit des distances, mais pas de direction ; une structure vectorielle définit des directions. Je n'explorerai pas plus avant les différences de structure entre formulations de la mécanique. Ce rapide examen me permet cependant de tirer plusieurs conclusions, par lesquelles je termine ce chapitre.

4.4 Les équations de la mécanique comme outils de représentation et d'inférence

En premier lieu, supposons que l'on adopte la thèse du réalisme structural, c'est-à-dire la forme du réalisme scientifique qui consiste à affirmer que les structures décrites par nos théories représentent (ou visent à représenter) les structures réelles du monde physique. C'est une option assez largement partagée par les philosophes des sciences, et qui correspond assez bien au type d'engagement réaliste dont font souvent preuve les scientifiques.⁷⁷ Elle est en effet considérée comme la meilleure manière de rendre compte du succès empirique de la science tout en évitant les objections classiques majeures au réalisme.⁷⁸ Dans une telle perspective, il semble bien que les différentes formulations de la mécanique ne disent pas la même chose de la structure du monde ; il n'en résulte aucune différence empirique, puisque le comportement des systèmes

⁷⁶Cela appelle une remarque à propos de la notion même de généralité, qui semble relative à l'usage que l'on fait d'une formulation. En effet, on vient de voir que la formulation hamiltonienne assigne une structure plus générale à l'espace des états que les formulations lagrangienne et newtonienne ; cependant, quand on se place dans le cadre de la théorie de la relativité, on affirme que la formulation lagrangienne est plus générale.

⁷⁷On attribue cette position à Poincaré, ainsi qu'à Duhem. Certains propos de Feynman (1965), quoique ambigus sur ce point, révèlent une certaine forme de réalisme structural.

⁷⁸Comme, par exemple, l'argument de la méta-induction pessimiste (voir Laudan, 1981).

physiques est le même quelle que soit la formulation adoptée, mais la structure géométrique sous-jacente du monde n'est pas la même. Ce constat conduit d'ailleurs Jill North (2009b, p. 26) à affirmer que, quand bien même cela n'implique pas la possibilité de les départager empiriquement, les trois formulations de la mécanique « ne sont même pas authentiquement équivalentes empiriquement. »⁷⁹

En outre, sans même s'engager en faveur du réalisme structural – ou de la version peut-être trop grossière que j'en décris ici⁸⁰ –, c'est-à-dire sans affirmer que la structure géométrique de l'espace dans lequel la mécanique décrit les systèmes vise à représenter la structure géométrique du monde, prendre conscience des contraintes géométriques associées à chaque interprétation permet d'explorer le contenu spécifiquement géométrique de la théorie. Si l'on conçoit la mécanique non pas seulement comme la science du mouvement dans l'espace et le temps, mais comme la théorie de l'espace et du temps en tant qu'ils contraignent les mouvements, cela prend toute son importance : les différentes formulations, en décrivant les phénomènes du mouvement à différents niveaux de généralité, n'en donnent pas la même image, et ne les expliquent pas de la même manière. Cela permet d'affirmer que l'identité des différentes formulations de la mécanique – la question de savoir si elles expriment, ou non, le même contenu – dépend en partie de ce que ses utilisateurs considèrent être son objet ; or, sur ce point, plusieurs positions sont défendables.

Je développerai ce dernier argument au chapitre 5 (section 3.3), consacré à un approfondissement de ma critique des approches visant à identifier formellement le contenu des théories. Pour l'heure, admettons que l'on peut légitimement répondre à l'argument qui précède en affirmant que la structure géométrique de l'espace sous-jacent à la description des mouvements est inessentielle au contenu de la mécanique, et que seul compte ce que les équations permettent de déduire quant au mouvement des systèmes. De ce point de vue, le contenu géométrique des différentes formulations ne fait pas partie du contenu de la théorie, qui n'affirme rien quant à la structure géométrique du monde ; il n'est que le cadre dans lequel les mouvements des systèmes sont *représentés*, cadre certes intéressant à étudier quand on cherche à analyser les relations des mathématiques et de la physique, mais qui ne nous apprend rien du monde. Outre qu'une telle réponse implique l'adoption d'une thèse philosophique sur la nature de la physique mathématique que l'on peut discuter – ce que je ne ferai pas –, elle repose, comme je vais le montrer, sur une conception insatisfaisante du

⁷⁹ Comme le remarque Olivier Darrigol (communication personnelle), cela montre que le réalisme structural n'est pas à définir en référence à la structure mathématique de l'appareil formel d'une théorie (ce que Poincaré, selon Olivier Darrigol, comprend très bien dans la mesure contestable où il est un réaliste structural).

⁸⁰ Voir note précédente.

contenu des théories.

En effet, en faisant abstraction de la manière dont les différentes formulations de la mécanique représentent les phénomènes – en ne prenant pas en compte la forme des équations, liée au choix d’un système de coordonnées – cette conception *dissocie tout à fait les fonctions représentative et inférentielle des théories*. J’ai montré, dans ce qui précède, que les différences pratiques liées à l’utilisation de l’une ou l’autre formulation sont le corollaire des différentes représentations mathématiques de l’état dynamique des systèmes. Ainsi, les différents types d’inférences que les agents doivent faire pour prédire et expliquer les phénomènes ont une contrepartie dans la structure même de l’espace géométrique dans lequel les différentes formes d’équations représentent les phénomènes.⁸¹ Refuser de prendre cela en considération dans une définition du contenu d’une théorie, c’est se condamner à ne pas comprendre comment, en exprimant des hypothèses à propos des phénomènes, cette théorie permet de les prédire et de les expliquer.

Il y a, certes, quelque chose comme un contenu de la mécanique, identifiable à l’ensemble de ses conséquences déductives *en principe*, puisque ses principes sont inter-déductibles. Cependant, affirmer que la mécanique est entièrement identifiable à ce contenu, abstraction faite de la manière dont elle exprime les hypothèses desquelles peut en principe être déduit cet « ensemble de conséquences » et, corrélativement, de la manière dont on y accède, revient à en faire une entité abstraite qui n’est ni outil de représentation ni outil d’inférence.

Un tel contenu désincarné n’est pas l’objet de l’activité théorique, qui consiste à construire, développer et manipuler des représentations pour tirer des inférences. Qu’il s’agisse de l’application des différents principes de la mécanique pour résoudre des problèmes concrets ou du développement même de la théorie – l’invention de nouvelles formulations et la découverte de nouvelles conséquences –, une telle activité est incompréhensible si l’on définit la théorie comme un tout, identifiable à un petit nombre de principes duquel est censé pouvoir être déduit tout son contenu.

En critiquant la notion de théorie qui sous-tend l’affirmation selon laquelle les formulations de la mécanique expriment une seule et même théorie, je ne prétends pas en donner une autre définition, qui permettrait de rendre compte de la multiplicité des formulations de la mécanique classique comme autant de théories. Cela reviendrait simplement à modifier les critères d’identification des théories et à déplacer les frontières qui les séparent les unes des autres. Or, en soulignant les différences entre

⁸¹Ce constat rejoint la manière dont Jeremy Butterfield (2004, p. 29) décrit les différences entre mécanique lagrangienne et mécanique newtonienne : la formulation lagrangienne fournit un schème pour résoudre des problèmes ; cela s’accompagne du fait qu’elle ne nous donne pas la même image du monde.

les formulations de la mécanique, je ne prétends pas affirmer qu'elles sont *des théories différentes*, mais que ces différences sont des *différences théoriques* qui méritent d'être prises en compte et étudiées en tant que telles.

Signalons d'ailleurs que le constat de ces différences entre les formulations de la mécanique, s'il remet sérieusement en question l'identité de son contenu, n'implique aucunement la nécessité d'opérer un *choix* entre ces formulations. Bien au contraire, cette multiplicité de formulations possibles, plus ou moins adaptées aux différents problèmes à résoudre, est une des forces de la théorie, puisqu'elle contribue non seulement à en augmenter l'efficacité prédictive, mais encore à faciliter des développements théoriques dans d'autres domaines de la physique. Le choix d'utiliser l'une ou l'autre formulation est dicté, en pratique, par le type de problème à résoudre ou par les généralisations que l'on cherche à opérer, mais il n'est en aucun cas question d'en choisir une comme étant la « meilleure », quel que soit le critère présidant à un tel jugement.⁸²

Plutôt que de proposer une autre définition de la notion de théorie, je suggère que cette notion, quand on cherche à l'identifier à un ensemble d'hypothèses, quel que soit le critère de cette identification, ne permet pas de rendre compte de la double dimension représentationnelle et inférentielle de l'activité théorique qui lui est pourtant essentielle. Il convient dès lors de se tourner vers d'autres unités d'analyse, et de développer d'autres notions qui permettent de rendre compte de cette double dimension.⁸³

Une telle approche devra en particulier prêter attention aux utilisateurs des théories, et à la forme des représentations qu'ils manipulent. On a vu, dans ce qui précède, à quel point la question de l'utilisation des formulations pour résoudre des problèmes et les différences de représentation mathématique qui les caractérisent sont liées. Une fois que l'on a pris acte de l'importance de la forme même des représentations utilisées, et que l'on renonce à faire des théories des « tous » abstraits dans ce que Paul Humphreys (2004, p. 77) appelle une perspective « sans propriétaire », le point de vue des agents devient primordial.

⁸²C'est une des raisons pour lesquelles Kuhn, comme on le verra au chapitre 6 (section 3.1.2, en particulier page 336), accorde peu d'attention au cas des différentes formulations de la mécanique classique : « elles ont coexisté sans être considérées incompatibles » (Kuhn, 1976, p. 198, note 10 ; 2000, p. 188, note 18).

⁸³Jeremy Butterfield (2004) suggère ainsi, dans son analyse de la mécanique lagrangienne, que les notions classiques de « lois » et de « modèles » ne sont pas suffisantes à une étude approfondie de la manière dont les théories permettent de résoudre des problèmes. La notion de « gabarit », proposée par Paul Humphreys (2004), vise aussi à permettre une analyse plus fine des procédures de calcul mises en œuvre dans l'application des théories ; j'exposerai ses propos sur la notion de gabarit au chapitre 6, section 3.2.

Pour finir, notons qu'une telle perspective invite, comme le suggère Paul Humphreys (2004, pp. 67-70), à repenser la cartographie même des disciplines scientifiques et des champs de l'activité théorique. En effet, « une même théorie » – traditionnellement identifiée à son domaine de phénomènes et à l'ensemble des choses qu'elle en dit *en principe* – peut être utilisée au moyen de différents types de représentation et, inversement, certains types de représentations se retrouvent dans d'autres domaines théoriques (le formalisme lagrangien en théorie des champs et en relativité générale, et le formalisme hamiltonien en mécanique quantique et statistique).⁸⁴ En parallèle d'une cartographie fondée sur le domaine des phénomènes étudiés et sur les conditions de vérité *en principe*, on peut envisager de tracer des frontières d'un autre type, qui tiendraient compte des types de représentations utilisés, et des raisonnements et pratiques qui leur sont associés.⁸⁵

Conclusion

Les trois formulations de la mécanique – newtonienne, lagrangienne et hamiltonienne – offrent une représentation différente des phénomènes du mouvement ; leurs hypothèses sont exprimées par des principes qui ne sont pas fondés sur les mêmes concepts, et qui sont formulés dans des langages mathématiques différents (calcul différentiel et calcul variationnel). La forme même des équations qui en découlent, exprimées dans des systèmes de coordonnées différents, a d'importantes conséquences sur la manière dont, en pratique, on les utilise pour résoudre des problèmes de mécanique.

J'ai montré, par une analyse du lien qui unit les différences de représentation mathématique aux différences pratiques liées à l'utilisation de l'une ou l'autre formulation, qu'une attention exclusive à la notion de théorie, identifiée à l'ensemble de ses conséquences déductives *en principe*, ne permet pas de rendre compte de la double

⁸⁴Dans une perspective toutefois différente de la mienne, Olivier Darrigol (2008) propose la notion de « module » pour rendre compte des similitudes inter-théoriques et des différences intra-théoriques dont les analyses classiques des théories et des relations inter-théoriques ne rendent pas compte. Il définit les modules comme des « composants essentiels » des théories « qui sont eux-mêmes des théories avec différents domaines d'application », et qui jouent un rôle fondamental dans l'application, la compréhension, la construction et la communication des théories. Son but, à la différence du mien, n'est pas tant d'insister sur la forme des représentations utilisées dans la pratique scientifique que sur la complexité des relations structurelles entre différentes théories (et modules). Il remet toutefois en question, comme je cherche à le faire, la pertinence des définitions classiques des théories et de l'unité (ou de la division) des sciences pour une analyse de l'activité théorique.

⁸⁵On retrouvera cette idée dans la deuxième partie de ce travail, au chapitre 6 (section 3.2.2). Elle sera au cœur de mon étude de cas sur la génétique, dans la troisième partie (chapitres 8 et 9).

dimension représentationnelle et inférentielle de l'activité théorique. Les différences mathématiques entre les formulations, qui sont objectives au sens où on peut les identifier sans faire référence aux opérations inférentielles des agents, mais qui n'affectent pas son contenu empirique *en principe*, ont une contrepartie pratique. Une fois l'importance de ces différences reconnues, on n'est plus justifié à ignorer l'utilisation des théories par les agents. J'ai finalement suggéré qu'une analyse de cette double dimension gagnerait à développer d'autres notions que celle de théorie, en adoptant la perspective des agents et en prêtant une attention particulière aux formes des représentations utilisées dans la pratique scientifique et aux raisonnements que leur utilisation met en jeu.

La suite de ce travail est destinée à élaborer et à mettre en œuvre cette approche. Le chapitre 3, au moyen de la notion de *version*, va approfondir les différences conceptuelles que l'on trouve *au sein* de la mécanique classique, d'un double point de vue diachronique et synchronique ; ce faisant il montrera, je l'espère, la fécondité d'une approche qui adopte la perspective des agents. Après un chapitre consacré à une critique des approches visant à définir formellement le contenu des théories (chapitre 5), les chapitres 6 et 7 seront consacrés à élaborer des outils permettant d'analyser l'interaction cognitive des agents avec les représentations qu'ils utilisent. La troisième partie de ce travail (chapitres 8 et 9) les mettra en œuvre dans une seconde étude de cas.

Chapitre 3

Architecture conceptuelle et versions de la mécanique classique



Louis-Joseph Lagrange. Gravure italienne du début du XIX^e siècle. Bibliothèque de l'Académie des Sciences.
Photo Jean-Loup Charmet.

Le chapitre 2 a montré que la notion de théorie, définie comme un ensemble d'hypothèses jointes à la totalité de leur conséquences déductives, ne permet pas de rendre compte de l'activité théorique, en tant qu'elle implique des agents utilisant différentes représentations pour tirer des inférences à propos des phénomènes qu'ils étudient. J'ai suggéré, en conséquence, que l'on gagnerait à prêter une attention plus grande aux agents, et aux représentations au moyen desquelles ils raisonnent quand ils apprennent, développent, et appliquent une théorie. Dans ce chapitre, je poursuis l'analyse de l'exemple de la mécanique classique en adoptant résolument la perspective des agents. Pour cela, j'introduis la notion de *version* : je propose d'appeler « version » la manière dont un agent, profane ou expert, comprend et utilise une théorie. Ce chapitre vise à préciser la définition de cette notion et à en montrer la fécondité. Ce faisant, j'aborderai plusieurs problèmes identifiés au chapitre 1.

– En premier lieu, j'approfondirai l'analyse des *différences conceptuelles* entre les formulations de la mécanique, que j'ai momentanément laissées de côté dans le chapitre 2. Les formulations de la mécanique, telles qu'on les enseigne aujourd'hui, se distinguent en effet non seulement par la forme des équations au moyen desquelles elles permettent de décrire le comportement des systèmes physiques (problème étudié au chapitre 2), mais aussi par les *concepts* qu'elles mettent en jeu et par l'*architecture des principes* au moyen desquels elles expriment leurs hypothèses. Ces différences, quoique gommées par l'identification du contenu de la mécanique à ses conséquences déductives *en principe*, sont toutefois objectives, au sens où elles sont identifiables sans qu'il soit nécessaire de faire référence aux agents qui utilisent la théorie. Un des buts de ce chapitre est cependant de montrer que la prise en compte des agents individuels – et de leurs versions – permet de mener une analyse plus fine de ces différences.

– Ce premier problème, *synchronique*, se double d'un second problème, *diachronique*. Comme je l'ai signalé au chapitre 1, les différentes formulations de la mécanique, aujourd'hui, sont le résultat d'un développement historique, que les catégories au moyen desquelles on étudie habituellement le changement scientifique (incommensurabilité ou réduction inter-théorique) ne permettent pas de décrire. En effet, contrairement à ce qui est le cas, par exemple, dans le passage de la mécanique newtonienne à la mécanique einsteinienne, les développements de la mécanique par Lagrange (1788) et Hamilton (1834), entre autres, ne consistent pas en une modification logique de la définition des concepts de Newton (1687), laquelle impliquerait une modification des conséquences déductives *en principe* de la théorie. Analyser les présentations successives de la mécanique comme l'expression de *versions* différentes me permettra de rendre compte de la dimension conceptuelle de ces changements.

– Enfin, l'introduction de la notion de version pour aborder ces deux ensembles de

problèmes permet de jeter un éclairage nouveau sur un troisième problème, d'ordre *historiographique*. Comme je l'ai signalé au chapitre 1, les formulations newtonienne, lagrangienne et hamiltonienne de la mécanique, telles qu'on les présente aujourd'hui, ne correspondent pas aux présentations de Newton (1687), Lagrange (1788) et Hamilton (1834). Pour autant, affirmer que les noms que l'on donne aux formulations d'aujourd'hui reposent sur une erreur de jugement historique reviendrait à réécrire une histoire tout aussi grossière que celle qui consiste à assimiler sans examen les travaux de Newton et de Lagrange aux formulations actuelles qui portent leurs noms. Comme je l'ai souligné au chapitre 1, évaluer les raisons qui permettent ou non d'attribuer, par exemple, le nom de Lagrange à la formulation lagrangienne d'aujourd'hui implique d'une part de préciser ce que l'on considère constituer le contenu de cette formulation et d'autre part d'identifier le contenu de la présentation de la mécanique proposée par Lagrange (1788). En particulier, cela implique de savoir si les différences (formelles et/ou conceptuelles) entre les formulations newtonienne et lagrangienne d'aujourd'hui sont le résultat des nouveautés (formelles et/ou conceptuelles) introduites par Lagrange. L'analyse de la version de Lagrange, telle qu'il l'exprime dans sa *Mécanique analytique* de 1788, suivie de celle de plusieurs autres versions de la mécanique exprimées par des physiciens et des philosophes postérieurs à Lagrange, me permettra de mettre en évidence que la distinction entre un développement formel et une avancée conceptuelle – au nom de laquelle on peut juger que Lagrange, pour ainsi dire, est ou n'est pas « lagrangien » – n'est pas tranchée. On verra en effet que le jugement que chacun porte sur la contribution historique de Lagrange *dépend de sa propre version de la mécanique*.

Dans ce qui suit, je reviens dans un premier temps (section 1) sur les différences conceptuelles entre les formulations de la mécanique telles qu'on les présente aujourd'hui. Ensuite (section 2), je définis la notion de version, comme une hypothèse utile pour rendre compte de l'activité théorique. La section 3 est consacrée à l'analyse de la version de la mécanique que Lagrange (1788) exprime dans la *Mécanique analytique* ; je prêterai une attention particulière au jugement que Lagrange porte lui-même sur l'avancée que constituent ses travaux. La section 4 présente les versions que plusieurs savants et philosophes postérieurs à Lagrange ont exprimées dans leurs présentations de la mécanique ; j'y montre en particulier le lien qui existe entre leurs versions et leur jugement sur la nature du développement historique de cette science. Enfin, la section 5 dresse le bilan de ces différentes versions et revient sur la différence entre avancée formelle et changement conceptuel.

1 L'architecture conceptuelle des formulations de la mécanique classique

Dans cette section, je souhaite montrer que, malgré l'inter-déductibilité de ses différents principes, il est un sens dans lequel on peut dire que la mécanique n'a pas la même signification selon le principe que l'on choisit de placer à son fondement. Dans cette perspective, ce qui distingue une mécanique de la force fondée sur un principe différentiel d'une mécanique de l'énergie fondée sur un principe variationnel est une différence d'ordre conceptuel, et non pas seulement formel.

1.1 La signification des principes dépend de leur place dans l'architecture déductive de la théorie

En premier lieu, selon le principe sur lequel on considère que repose l'édifice tout entier de la mécanique et dont les autres principes sont des conséquences déductives, les concepts premiers de la théorie, c'est-à-dire ceux qui ne sont pas définis par elle mais qui lui donnent sa signification physique, ne sont pas les mêmes. Ainsi, si l'on place le principe de Hamilton au fondement de la théorie, le concept premier est celui d'énergie. La mécanique, en tant qu'elle repose sur ce principe, *dit* qu'une certaine quantité, qui a la dimension d'une énergie, est stationnaire au cours du mouvement. Dans cette perspective, l'expression de la deuxième loi de Newton que l'on « retrouve » à partir du principe de Hamilton peut être considérée comme une formule qui tire (au moins partiellement) sa signification d'une affirmation portant sur l'énergie des systèmes.

En revanche, quand elle est placée au fondement de la théorie, la deuxième loi de Newton affirme quelque chose à propos du monde, à savoir que la force et l'accélération sont proportionnelles. Ce que dit la théorie newtonienne, c'est cela. En conséquence, la théorie fondée sur un tel principe tire sa signification de celles des concepts de masse et de force. À son tour, dans ce cadre, le principe de moindre action, déduit de cette loi, n'est qu'une formule pour ainsi dire aveugle n'apportant pas de signification nouvelle à la théorie.

En outre, selon que la mécanique est présentée comme reposant sur le principe de Newton ou sur celui de Hamilton, l'architecture déductive globale de la théorie change d'aspect. En effet, les liens déductifs entre les différents principes ne sont pas les mêmes. Ils le sont, certes, en principe, puisque l'on peut tous les déduire les uns des autres, mais le chemin à suivre n'est pas le même, au sens où l'*ordre* déductif dans lequel ils sont présentés est différent ; la place de chacun des principes dans l'édifice n'est pas la même.

En plaçant la deuxième loi de Newton au fondement de la théorie, on fait du principe de Hamilton une conséquence « périphérique » des principes fondamentaux de la mécanique. En revanche, en plaçant le principe de Hamilton au fondement de la théorie, on met en évidence l'équivalence des équations du mouvement déduites d'un principe différentiel avec un principe variationnel formulé en termes d'énergie ; ce faisant, on établit de nouveaux liens déductifs entre les différents principes, ainsi qu'entre la mécanique et les autres théories faisant usage du concept d'énergie.

Si l'on considère qu'une telle modification dans l'architecture de la théorie est purement formelle, et que l'expression de la loi de Newton a le même contenu quels que soient les liens établis entre elle et les autres principes, alors on doit admettre que toutes les conséquences déductives que le principe de Hamilton permet d'atteindre *étaient déjà* des conséquences du principe fondamental de Newton avant que celui de Hamilton soit formulé. Dans une telle perspective, tout ce que l'on peut, en principe, déduire du principe de Newton fait partie de son contenu.

C'est parce que de telles affirmations me semblent hautement insatisfaisantes que je propose de considérer que les différentes manières de présenter l'architecture déductive de la théorie correspondent à des différences conceptuelles, et non pas simplement formelles. Autrement dit, dans ma perspective, le contenu des différents principes – et des concepts qu'ils emploient – est affecté par la modification de leur place dans l'ordre déductif de la théorie et par l'introduction de nouveaux principes, quand bien même ces derniers sont obtenus par une opération purement déductive.

1.2 Différence architectonique et différence logique

Pour préciser ce que j'entends ici par « différence conceptuelle », comparons un instant ce cas avec celui de la différence entre mécaniques newtonienne et einsteinienne, dans l'analyse que Kuhn (1962/1970, pp. 101-102) en donne.¹ Kuhn affirme que l'expression « $\mathbf{F} = m\mathbf{a}$ » que l'on peut déduire, moyennant une approximation, de la théorie d'Einstein n'est pas la loi de Newton, bien qu'elle en ait la même forme, parce que dans le cadre einsteinien la définition de la masse est différente de celle qu'elle a dans le cadre newtonien. Affirmer, comme je le fais, que la loi de Newton n'a pas exactement le même contenu selon qu'elle est au principe de la théorie ou qu'on la « retrouve » à partir du principe de Hamilton peut paraître reposer sur des pré-supposés analogues à ceux de Kuhn. Il existe pourtant une différence fondamentale

¹ Je présenterai cette analyse de façon plus détaillée au chapitre 6 (section 1.2.2). Dans ce même chapitre (section 1.3), j'approfondirai la comparaison de la conception kuhnienne de la signification et du changement conceptuel qui le conduit à soutenir la thèse de l'incommensurabilité des paradigmes avec l'approche que je défends au moyen de la notion de version.

entre ces deux thèses.

Si Kuhn refuse d'appeler « loi de Newton » l'expression « $\mathbf{F} = m\mathbf{a}$ » déduite de la théorie d'Einstein, c'est parce que, comme je viens de le dire, il existe une différence entre les *définitions* newtonienne et einsteinienne de la masse. L'incommensurabilité des paradigmes newtonien et einsteinien, en raison de laquelle cette même expression n'a pas la même signification dans les deux théories, repose sur une *différence logique* entre leurs définitions du concept de masse, et non pas sur une différence dans l'*ordre déductif* des principes. Le holisme sémantique qui sous-tend l'argument de Kuhn, selon lequel les concepts d'une théorie tirent leur signification de l'ensemble de la théorie, n'est pas un holisme *situé*, si l'on entend par là une position – comme celle que j'adopte – qui consiste à considérer qu'un principe tire sa signification non pas seulement du système théorique auquel il appartient, mais aussi de la *place* particulière qu'il y occupe.

Mon propos est de montrer qu'une différence *architectonique*, même en l'absence de différence logique, implique une variation conceptuelle. En revanche, alors que le holisme sémantique qui sous-tend l'argument de Kuhn le conduit à affirmer une différence de signification *totale* entre l'expression de la loi de Newton dans le cadre newtonien et la même expression dans le cadre einsteinien – au point que le symbole représentant le concept de masse n'a pas le même référent dans les deux paradigmes (Kuhn, 1962/1970, p. 101) –, je ne considère en aucun cas les différences conceptuelles entre les formulations de la mécanique comme des exemples d'incommensurabilité. Ce que je rejette, en refusant une conception logico-empirique du contenu des théories, c'est aussi l'idée selon laquelle les différences conceptuelles sont affaire de tout ou rien. Je souhaite montrer qu'au sein d'un même édifice logique, c'est-à-dire d'une même « théorie », il existe des variations authentiquement conceptuelles.

1.3 Conception vériconditionnelle et conception inférentielle de la signification

L'affirmation selon laquelle les différences architectoniques entre les formulations de la mécanique sont des différences conceptuelles implique l'adoption d'une conception de la signification qui se distingue nettement d'une conception vériconditionnelle, selon laquelle deux ensembles d'énoncés ont la même signification si et seulement s'ils ont les mêmes conditions de vérité. La notion de signification qui sous-tend mon analyse accorde, contrairement à une approche en termes de conditions de vérité, une importance à l'*usage* que l'on en fait dans des processus inférentiels. Autrement dit, plutôt que de considérer que la signification d'un principe – et des concepts qui y figurent – lui vient de l'ensemble de ses conséquences déductives en principe, je sug-

gère qu'elle est aussi déterminée par les inférences au moyen desquelles on l'obtient, ainsi que par les inférences qu'il facilite, en pratique.

Ici la distinction entre l'analyse de la connaissance scientifique *en principe* et *en pratique* prend un sens légèrement différent de celle qui a guidé l'analyse du chapitre 2. Dans les deux cas, il s'agit d'introduire, dans l'analyse du contenu de la mécanique, le point de vue des agents. Dans l'analyse du chapitre 2, j'ai cherché à souligner, en adoptant ce point de vue, la différence entre ce que l'on peut en principe déduire d'une équation, et ce qu'il est possible, en pratique, d'en faire. Dans l'analyse du présent chapitre, j'oppose à une conception synoptique de la structure logique d'une théorie une conception qui prend en compte le fait que son utilisation se fait toujours dans une certaine perspective : on n'accède pas à la théorie comme à un « tout », mais on y accède plutôt par tel ou tel « bout », que l'on prend comme principe premier, et duquel on déduit les autres. Autrement dit, je prends en compte le caractère situé de l'accès que les agents ont à la théorie. Dans cette perspective, l'ordre dans lequel les principes sont déduits les uns des autres, et la hiérarchie que l'on établit entre eux, est de première importance pour en établir le contenu.

Les différences que je viens de souligner dans l'ordre déductif des différents principes, que l'on décide ou non de les appeler « conceptuelles », sont toutefois des différences objectives, au sens où j'ai pu les mettre en évidence sans faire mention de l'usage qu'en ferait un agent particulier. Pour autant, elles ont des conséquences sur les inférences que les agents doivent et peuvent effectivement faire. L'hypothèse des versions, que je propose dans la prochaine section, est destinée à aborder ce type de différence sous l'angle non pas seulement de l'ordre déductif objectif, mais sous celui des raisonnements que les agents conduisent effectivement quand ils utilisent la mécanique. C'est sous cet angle, plus « subjectif », que j'entends étudier les différences que j'appelle « conceptuelles » entre les formulations de la mécanique.

2 La notion de version

Dans cette section, j'introduis la notion de version, que je propose pour formuler une hypothèse permettant de rendre compte de plusieurs activités que j'appelle *théoriques*. Ces activités incluent aussi bien l'apprentissage que l'application et le développement d'un ensemble d'hypothèses théoriques à propos d'un domaine de phénomènes. Pour plus de simplicité, je désignerai cet ensemble d'hypothèses théoriques par le terme « théorie », sans que cela implique l'existence d'un critère d'identification précis des théories.

Intuitivement, la notion de version correspond à *la manière dont un sujet comprend une théorie*, c'est-à-dire la perspective particulière depuis laquelle il accède à

son contenu. Autrement dit, et tout aussi informellement, c'est la théorie « de » ou « pour » l'agent, la théorie *en tant qu'elle est comprise* par l'agent.

L'essentiel de la présente section est consacré à approfondir cette définition de la notion de version. Après avoir esquissé quelques pistes que l'hypothèse des versions permet d'emprunter pour étudier l'activité théorique, j'examine la relation entre les versions ainsi définies et les formulations de la mécanique, puis je présente la méthode adoptée dans la suite du chapitre pour étudier les versions exprimées par certains acteurs de l'histoire de la mécanique.

2.1 La notion de version comme hypothèse permettant de rendre compte de l'activité théorique

Je suppose que tout agent confronté à une théorie, du débutant à qui elle est enseignée à l'expert qui la développe ou la convoque pour prédire et expliquer des phénomènes, se forme un ensemble de représentations mentales, liées à l'utilisation de cette théorie.² Cet ensemble de représentations mentales est sa *version* de cette théorie.

Pour emprunter le vocabulaire de Hertz (1894), qui affirme que « le procédé que nous utilisons invariablement pour dériver le futur à partir du passé et parvenir ainsi à la prévision souhaitée » consiste à « fabriquer en nous des simulacres, ou des symboles internes des objets externes » (Hertz, 1894, p. 67) – simulacres qu'il appelle aussi « images » –, ma version de la mécanique, c'est l'ensemble des « images » ou représentations mentales que je me forme au cours de mon apprentissage et de mon utilisation de cette théorie. C'est en manipulant ces représentations, le plus souvent au moyen d'un support externe (par exemple, en écrivant des équations), que je parviens à tirer les inférences permettant de prédire et d'expliquer les phénomènes ainsi représentés. Étudier l'activité théorique au moyen de la notion de version revient donc, plutôt que de chercher à reconstruire le contenu « objectif » d'un ensemble d'hypothèses, indépendamment de la compréhension qu'en ont les agents, à analyser ce que signifie cet ensemble d'hypothèses pour l'agent qui les utilise.

Avant de poursuivre, il convient d'écarter une ambiguïté : dire que ma version de la mécanique consiste en un ensemble de représentations mentales n'implique pas qu'elle corresponde simplement à « ma vision des choses », c'est-à-dire à la façon dont je me représente les phénomènes du mouvement, quelle qu'elle soit. Ma version de la mécanique se forme au cours de mon apprentissage de la mécanique, et elle

²Dans ce qui suit, j'entendrai le plus souvent « utiliser une théorie » en un sens large, servant à désigner toute forme d'activité qui consiste à construire et à développer des hypothèses théoriques, à les appliquer à la résolution de problèmes particuliers, mais aussi à les enseigner et à les apprendre.

est contrainte par les règles d'utilisation des représentations (externes) servant à l'exprimer. Un usage fantaisiste de l'équation exprimant la deuxième loi de Newton n'émane donc pas d'une version de la mécanique newtonienne.

La version de chacun se caractérise par une certaine manière d'ordonner les différents principes de la mécanique, de les relier au monde empirique, et par la signification qu'il attribue aux concepts qui y figurent. Elle se manifeste par l'usage qu'il en fait pour représenter les phénomènes et pour raisonner à leur propos. Elle dépend à la fois de l'apprentissage qu'il a suivi, de ses connaissances d'arrière-plan, de ses savoir-faire, et plus largement de tout un ensemble de facteurs que l'on peut appeler « psychologiques » : ses habitudes de raisonnement, ses options métaphysiques implicites ou explicites, ses préférences et ses engagements théoriques.

Ici aussi, une mise en garde s'impose. Dire que certaines particularités psychologiques des agents déterminent (partiellement) leur version de la mécanique ne veut pas dire que ma version de la mécanique est affectée par toutes les représentations mentales qui naissent chez moi par association d'idées avec, par exemple, le terme « force », ou même par le fait que j'aie intérêt à résoudre un problème de mécanique pour réussir un examen. Aussi, les facteurs que j'appelle ici « psychologiques » n'incluent pas tout ce qui constitue la psychologie des agents au sens large. Ils ne se réduisent cependant pas aux seules *capacités* cognitives, dans la mesure où, comme je m'efforcerai de le montrer, les préférences et les intérêts théoriques des chercheurs jouent aussi un rôle dans leur utilisation et leur compréhension d'un ensemble d'hypothèses.

En outre, en supposant que les versions des agents sont partiellement déterminées par ces facteurs, mon but n'est pas d'étudier *la manière dont ils la déterminent* ; comme cela apparaîtra dans la suite de ce travail, c'est aux versions elles-mêmes, et plus précisément à la façon dont elles se manifestent dans la pratique des agents, que je vais m'intéresser. Autrement dit, à supposer même que certaines associations d'idées comme celles mentionnées ci-dessus aient une incidence sur l'usage qu'un agent fait des équations de la mécanique, ce qui va m'intéresser dans ce travail est *l'usage* qu'il en fait, et non les déterminants psychologiques de cet usage. On peut supposer que deux agents, pour des raisons psychologiques différentes, ont la même version – font le même usage – de la mécanique. Le fait que cette version soit le fruit de deux ensembles de déterminants psychologiques différents n'a pas de pertinence pour l'analyse que j'entends mener. Cependant, décrire les versions des agents comme un ensemble de représentations mentales implique de les considérer comme un élément de leur vie psychologique. La suite de cette section, en développant cette caractérisation sommaire, va me permettre de préciser cela et de montrer comment l'hypothèse des versions permet de décrire différentes activités théoriques, depuis l'apprentissage

d'une théorie par le débutant jusqu'à son développement et son approfondissement par l'expert.

2.1.1 Apprendre la mécanique : développer et enrichir sa version

Prenons un débutant : sa version minimale lui permet d'utiliser la deuxième loi de Newton dans des cas très simples, sans qu'il soit en mesure d'en tirer des explications pour des cas plus compliqués – dans lesquels il ne reconnaît peut-être même pas des situations susceptibles d'être décrites au moyen de cette loi – ni de voir le lien déductif qui la relie aux théorèmes de conservation, au principe de d'Alembert et encore moins aux équations de Lagrange et au principe de Hamilton. En outre, ses compétences restreintes en mathématiques limitent grandement sa capacité à tirer des inférences à partir des principes qu'il connaît. Sa version n'est pas à strictement parler fautive – dès lors qu'il utilise correctement la deuxième loi de Newton dans des cas simples – mais elle est très limitée.

On peut alors décrire le processus d'apprentissage de la mécanique comme celui de l'enrichissement et du développement de cette version minimale. Ce processus consiste à la fois à tisser de nouveaux liens inférentiels entre différentes représentations – par exemple, la deuxième loi de Newton et les théorèmes de conservation –, et à apprendre à les relier aux phénomènes empiriques dans différentes situations. L'étudiant apprend ainsi progressivement à voir l'unité des différents phénomènes du mouvement, derrière leur diversité apparente, et à représenter des situations de plus en plus diverses et de plus en plus compliquées au moyen des lois de la mécanique. L'apprentissage consiste aussi à rendre ces différents liens inférentiels plus « solides », au sens où l'étudiant acquiert progressivement une plus grande maîtrise du calcul différentiel et d'autres règles de déduction mathématique qui lui permettent d'effectuer des inférences plus rapides. On peut donc décrire l'apprentissage de la mécanique par un agent comme un processus par lequel sa version devient plus complète, plus développée et plus robuste.

En outre, une version de la mécanique, en tant qu'ensemble de représentations mentales, n'est pas indépendante des autres représentations mentales de l'agent. Dans la mesure où les termes utilisés pour exprimer les concepts de la mécanique sont pour la plupart empruntés au langage ordinaire, il est fort probable que le sens courant du mot « force », par exemple, joue un rôle non négligeable dans la compréhension que le débutant a de la mécanique et l'aide initialement à identifier les forces en présence dans une situation particulière. Apprendre la mécanique consiste cependant aussi à modifier progressivement – ou en tous cas à enrichir – l'usage que l'on fait de ces termes, et à rendre sa version plus indépendante des représentations pré-scientifiques

du mouvement.

En conséquence, alors que le novice ne fait d'abord pas de distinction très nette entre le concept non scientifique de force et celui qui figure dans la deuxième loi de Newton, l'expert a une version de la mécanique qui est moins empreinte de la signification courante des termes. Alors que la version du débutant, encore mal « détachée » de ses représentations non scientifiques et trop fragile pour lui permettre de passer aisément des équations newtoniennes aux équations lagrangiennes, ne lui permet pas de voir l'équivalence entre une explication en termes de force et une explication en termes d'énergie, on peut supposer que l'expert, familier des équations différentielles et du calcul variationnel, ne « voit » même plus la différence entre les formulations de la mécanique, tant il passe de l'une à l'autre sans difficulté. Sa perspective sur la mécanique lui en donne ainsi une vue d'ensemble plus complète que celle d'un profane.

2.1.2 Les versions des experts

La notion de version ne permet pas seulement de faire une distinction entre compréhension profane et compréhension experte. Décrire l'expertise comme la possession d'une version relativement robuste et complète de la mécanique ne signifie pas pour autant que les versions des différents experts soient identiques entre elles. On peut en effet supposer que, selon le domaine de spécialité de chacun, les types de raisonnement auxquels il est habitué, ainsi que ses engagements théoriques, les versions diffèrent les unes des autres.

En effet, les processus inférentiels mis en œuvre par les agents au cours de leurs raisonnements, quoique fortement contraints par les relations logiques entre les concepts de la théorie et les règles de déduction mathématique en vertu desquelles une inférence est correcte ou fautive, sont sous-déterminés par ces relations et ces règles. Autrement dit, même en les respectant, on peut ordonner les différents principes de la théorie de plusieurs manières différentes, et il reste un grand nombre de « chemins » inférentiels possibles.

Par exemple, quand bien même chaque expert a une version de la mécanique qui lui permet de passer aisément d'une formulation à l'autre et de voir l'équivalence des descriptions en termes de force et en termes d'énergie, on peut supposer qu'un spécialiste de physique quantique a acquis une telle familiarité avec les équations hamiltoniennes que c'est sous ce formalisme qu'il utilise le plus naturellement les lois de la mécanique et qu'il préfère, autant que possible, se représenter les phénomènes du mouvement au moyen de ces équations.

De plus, quoique l'expertise s'accompagne d'une distinction nette entre le sens

ordinaire et le sens scientifique des concepts (comme ceux de force et d'énergie), on peut raisonnablement supposer que leur usage ordinaire, ainsi que le rôle qu'ils jouent dans d'autres théories scientifiques, a une influence sur la manière dont on les utilise pour représenter les phénomènes.³ En conséquence, il est probable que les autres connaissances théoriques, mais également les options métaphysiques, implicites ou explicites, d'un physicien orientent son regard sur les lois de la mécanique et la signification qu'il leur donne : certains préféreront se représenter les phénomènes du mouvement en termes d'énergie, et ne voir dans la notion de force qu'un symbole vide de signification ; d'autres, d'esprit plus « mécaniste », préféreront penser qu'il y a effectivement des forces dans le monde, et que le concept d'énergie est d'utilisation plus pratique dans certains cas mais tire sa signification de sa définition en termes de force.

On peut également supposer qu'il existe plusieurs « types d'esprits », certains préférant manipuler des symboles mathématiques sans considération de la signification courante des concepts qu'ils désignent, d'autres accordant une plus grande part, dans leurs raisonnements, à ce que l'on appelle couramment l'« intuition physique » ou encore « l'imagination ». C'est ce que suggère Maxwell, dont les réflexions sur l'importance heuristique des images et des analogies dans le raisonnement scientifique ont fait l'objet de plusieurs travaux récents en philosophie des sciences⁴ :

[...] certains esprits se satisfont de la contemplation de pures quantités présentées à leurs yeux par des symboles, et à leurs esprits sous une forme que seuls les mathématiciens peuvent concevoir.

D'autres suivront plus volontiers des formes géométriques, qu'ils tracent sur le papier, ou construisent dans l'espace devant eux.

D'autres encore ne sont satisfaits que lorsqu'ils sont en mesure de projeter toute

³L'importance du sens courant des mots utilisés pour exprimer des concepts scientifiques, et des analogies et métaphores qu'ils suggèrent, déjà soulignée par plusieurs physiciens du dix-neuvième siècle, et en particulier par Maxwell (1890), fait l'objet d'un regain d'intérêt de la part de plusieurs philosophes des sciences aujourd'hui. Ainsi, par exemple, Evelyn Fox Keller (1983, 1985, 1989, 1995) montre, au moyen d'exemples tirés de la génétique et de la biologie évolutionniste, que certaines métaphores (comme celles suggérées par les expressions « gène égoïste », « action génique », « compétition », et tout particulièrement la métaphore de l'information en biologie moléculaire) déterminent partiellement les questions que les scientifiques cherchent à résoudre et les types d'explications qu'ils jugent satisfaisantes. L'idée selon laquelle la compréhension du langage en général – et non pas seulement du langage scientifique – est dirigée par un certain nombre de métaphores ancrées dans la manière dont nous nous représentons le monde (par exemple, les métaphores spatiales du « haut » et du « bas », omniprésentes dans notre langage, comme en témoignent toutes les expressions évoquant la « hausse » ou la « baisse » d'une valeur) a été développée dans le programme de recherche en linguistique cognitive de Georg Lakoff et Mark Johnson (voir par exemple Lakoff et Johnson, 1980).

⁴Voir par exemple (Nersessian, 1984, 1992a, 2002b; Cat, 2001).

leur énergie physique dans la scène qu'ils évoquent. Ils apprennent à quelle vitesse les planètes se déplacent dans l'espace, et ils en tirent un délicieux sentiment d'ivresse. Ils calculent les forces que les corps célestes exercent les uns sur les autres, et ils sentent leurs propres muscles contractés par l'effort.

Pour de tels hommes, quantité de mouvement, énergie, masse, ne sont pas que les expressions abstraites des résultats de l'enquête scientifique. Ils sont des mots puissants, qui remuent leur âme comme des souvenirs d'enfance. (Maxwell, 1890, vol. 2, p. 220)

Quoique ces différents « types d'esprits » utilisent tous les équations de la mécanique en suivant strictement les règles de déduction mathématique – il n'est pas question ici du poète qui tirera de la deuxième loi de Newton des conséquences non scientifiques sur l'ordre du monde –, il est fort probable que les représentations mentales au moyen desquelles ils raisonnent ne soient pas précisément les mêmes.

Ainsi, celui qui « préfère » les démonstrations géométriques ne verra dans l'utilisation du calcul différentiel qu'un expédient permettant de tirer des inférences plus rapides, mais se formera autant que possible des représentations géométriques des phénomènes ainsi décrits. De même, tout en maîtrisant le calcul variationnel et en sachant utiliser les équations lagrangiennes et hamiltoniennes, le physicien chez qui le concept de force, loin d'être un symbole abstrait, « remue l'âme comme un souvenir d'enfance » pourra ainsi considérer que les formulations analytiques sont des réécritures commodées de la mécanique, qui, pour autant, tire sa signification physique du concept de force.

En conséquence, la version de chacun de ces individus sera légèrement différente de celle des autres : celui qui se représente plus volontiers les phénomènes du mouvement au moyen du concept de force considérera que le principe fondamental de la théorie est bien celui de Newton, et que les autres en sont des conséquences. Le spécialiste de physique quantique, au contraire, placera au fondement de la théorie le principe de Hamilton, et, dans sa version, les lois de la mécanique classique seront très fortement liées à celles de la mécanique quantique.

Ainsi définie, l'hypothèse des versions semble permettre non seulement de rendre compte de l'apprentissage, mais aussi du développement d'une théorie et de l'exploration du contenu de chacun de ses principes, qui conduit parfois à la formulation de conséquences nouvelles : en effet, de même que l'on peut décrire le processus par lequel l'étudiant apprend la mécanique comme l'élargissement de sa version, on peut décrire les raisonnements qui ont conduit, par exemple, Hamilton à introduire de nouveaux principes et à réorganiser les principes existants, comme le développement et l'approfondissement de la sienne.⁵ En conséquence, on peut supposer que, selon

⁵Voir (Barberousse, 2008) pour une analyse du changement conceptuel à l'œuvre dans le cas de

la version qu'ils en ont, les différents experts ne sont pas portés à développer une théorie de la même manière. Cette supposition, ainsi que les précédentes, trouveront un écho dans l'analyse des versions de certains acteurs de l'histoire de la mécanique proposée aux sections 3 et 4 du présent chapitre.

En supposant que tout agent, du débutant à l'expert, se forme ainsi une version de la mécanique, élaborée et mobilisée aussi bien au cours de l'apprentissage que du développement et de l'approfondissement d'une théorie, je ne prétends en aucun cas nier qu'il existe une différence importante entre ma version fragile et partielle de la mécanique et celle, par exemple, de Hamilton, qui, en plus d'être plus robuste et complète, est *innovante*, au sens où elle donne lieu à une manière inédite de formuler et de relier les principes et les équations de la mécanique entre eux. Cependant, décrire l'apprentissage, le développement et l'invention d'hypothèses théoriques au moyen de la notion de version permet de supposer que, malgré d'importantes différences, il y a quelque chose de commun à ces différentes activités, que je propose de considérer comme des *activités théoriques*.

2.2 Comment étudier les versions de la mécanique ?

Jusqu'ici, la notion de version a été introduite à titre d'hypothèse. La manière dont cette dernière est déterminée par différents paramètres – l'enseignement suivi par l'agent, les autres connaissances dont il dispose, ses préférences et talents personnels, etc. – n'est pas, en tant que telle, accessible au philosophe et à l'historien, et n'a été, dans ce qui précède, que l'objet de suppositions. Les versions elles-mêmes, en tant que représentations mentales, ne peuvent être étudiées que par la manière dont elles se manifestent.

La première voie par laquelle se manifeste – et se construit – la version d'un individu est sa *pratique* de la théorie, c'est-à-dire la manière dont il utilise ses différentes hypothèses pour résoudre des problèmes particuliers et pour répondre à des demandes d'explication.⁶ Cette pratique se caractérise souvent par la construction

la mécanique à l'aide de la notion d'*explicitation* [*making explicit*] empruntée à Robert Brandom (2000).

⁶On peut décrire les expériences de psychologie qui portent sur la manière dont les agents, experts ou profanes, résolvent des problèmes de mécanique comme des enquêtes visant à expliciter leurs versions de cette théorie (voir par exemple Larkin *et al.*, 1980; White, 1983; Larkin, 1983). Dans certains cas, ces enquêtes portent sur ce que l'on appelle la « physique naïve », c'est-à-dire la manière dont les sujets ignorants de la mécanique – la plupart du temps, de jeunes enfants – se représentent les phénomènes du mouvement. On ne peut pas, dans ce cas, parler de versions *de la mécanique*. Cependant, certaines expériences portent aussi sur la compréhension que des sujets profanes ont des principes de la mécanique qui leur sont enseignés et de la manière dont ils les utilisent, une fois confrontés à des problèmes. Ce qu'ils mettent en œuvre alors est bien une certaine

et la manipulation de représentations *externes* au moyen desquelles l'agent tire des inférences à propos des phénomènes représentés.

En effet, introduire l'hypothèse des versions – c'est-à-dire l'hypothèse selon laquelle chaque utilisateur de la mécanique se forme un ensemble de représentations mentales associées à l'utilisation de cette théorie – ne revient pas à nier l'importance, pour les processus inférentiels des agents, des représentations externes. Décrire les raisonnements et les calculs à l'œuvre dans l'activité théorique comme la construction et la manipulation d'un ensemble de représentations mentales ne signifie pas que ces processus soient indépendants du travail d'écriture et de manipulation de représentations externes comme des équations ou des graphes. Autrement dit, je ne veux en aucun cas suggérer l'idée naïve selon laquelle les agents, qu'il s'agisse du débutant qui apprend ou de l'expert qui développe et invente, aurait « dans la tête » une certaine représentation, qu'il pourrait le cas échéant exprimer, après coup, à l'aide de traces écrites concrètes. Bien au contraire, comme il apparaîtra dans la suite de mon travail, la méthode que j'y mets en œuvre consiste à centrer mon analyse sur les représentations externes construites et manipulées par les agents.⁷

À la limite, comme on vient de le voir, tout individu, qu'il s'agisse de Newton, de Lagrange, de Hamilton, de Hertz, de Mach, de l'auteur d'un manuel de mécanique, d'un géomètre non physicien, d'un spécialiste de la théorie quantique, d'un enseignant, d'un ingénieur, d'un étudiant ou encore d'un philosophe des sciences, a sa propre version de la mécanique. Elle est, selon les cas, plus ou moins confuse, développée, complète et robuste. La mécanique – comme n'importe quelle théorie – fait donc l'objet d'autant de versions qu'il y a d'individus qui l'utilisent et la comprennent, partiellement ou totalement.

Cependant, parmi cette multitude de versions possibles, la plupart sont des perspectives plus ou moins complètes sur un corps de connaissances qui a déjà été établi et présenté de manière explicite : quand j'étudie la mécanique, j'apprends à me servir de représentations externes que je trouve dans des manuels, et je m'efforce de suivre les démonstrations et les déductions qui ont été faites par d'autres que moi. Dans

version de la mécanique, certes très pauvre et parfois incohérente, parce qu'entâchée de croyances fausses.

⁷L'interaction entre les représentations mentales des agents et les représentations externes qu'ils utilisent est sans doute beaucoup plus complexe que l'image naïve mentionnée ci-dessus ne le laisse entendre. De nombreux travaux en sciences cognitives, en particulier ceux qui portent sur la cognition distribuée (voir par exemple Hutchins, 1995; Clark, 1997), cherchent aujourd'hui à approfondir l'analyse de la relation entre représentations mentales et représentations externes, supposant en particulier que ces dernières ne sont pas de simples supports pour le raisonnement que les agents opèrent au moyen de représentations mentales, mais sont constitutives de ce raisonnement, qui se ferait à la fois « dans » et « hors de la tête » des agents.

certains cas, les agents – théoriciens, philosophes, mais aussi auteurs de manuels – rendent leur version *explicite* au moyen d’une *présentation* publique de la mécanique. C’est là la seconde voie par laquelle les versions de certains agents sont accessibles à l’analyste, et c’est celle que j’entends exploiter dans la suite de ce chapitre : je me propose en effet de considérer les traces écrites des travaux des savants qui ont contribué à l’histoire de la mécanique, qu’il s’agisse de la *Mécanique analytique* de Lagrange (1788), des articles de Hamilton (1834), de l’« exposé historique et critique » du développement de la mécanique par Mach (1883), ou encore des *Principes de la mécanique* de Hertz (1894), comme une expression (au moins partielle) de leur version – et le seul accès que nous y ayons.

On retrouvera, au cours de l’analyse des sections 3 et 4, le lien entre les versions de ces individus et les différents déterminants que j’ai mentionnés dans la présente section. Auparavant, il convient de dire un mot de la manière dont la notion de version permet d’éclairer les différences conceptuelles entre les formulations de la mécanique, pointées à la section 1.

2.3 Versions et formulations

J’ai proposé, à la section 1 du présent chapitre, de considérer les différences architectoniques entre les formulations actuelles de la mécanique comme d’authentiques différences conceptuelles, et non pas seulement comme des différences formelles. Cela revient à dire que les différents principes et concepts de la mécanique n’ont pas la même signification selon la formulation que l’on utilise. Cela implique, comme je l’ai suggéré, l’adoption d’une conception de la signification qui prête attention au rôle que les différentes représentations jouent dans les processus inférentiels des agents, et non pas seulement à leurs conditions de vérité.

C’est – entre autres – dans ce but que j’ai introduit la notion de version : on peut en effet considérer les différentes formulations de la mécanique – qui ont les mêmes conditions de vérité *en principe* – comme l’expression d’autant de versions de cette théorie. En effet, bien que, comme je l’ai souligné au chapitre 2 (page 81), il ne soit pas question de choisir entre les différentes formulations qui sont, selon les cas, plus ou moins adaptées à la résolution de problèmes, on peut, en un autre sens, les considérer comme des manières concurrentes de comprendre et d’utiliser la mécanique.

L’étudiant n’est certes pas sommé de choisir entre formulation newtonienne et formulations analytiques, et doit au contraire développer sa capacité à utiliser l’un et l’autre formalisme. Il n’en reste pas moins que, commençant par apprendre la mécanique dans la perspective newtonienne, le débutant conservera probablement une version newtonienne de la mécanique jusqu’à devenir familier des équations lagran-

giennes ou hamiltoniennes au point de voir tout le contenu de la mécanique depuis cette nouvelle perspective.

De même, l'expertise se caractérise entre autres par la maîtrise des différentes formulations et par la capacité à passer aisément de l'une à l'autre – et donc par la possession d'une version plus « globale » de la mécanique, qui en inclut les différentes formulations possibles. Cependant, considérer les différentes formulations comme les expressions de versions différentes permet de rendre compte à la fois de l'intuition selon laquelle elles ne disent pas exactement la même chose des phénomènes et du fait qu'elles ne permettent pas de déduire les mêmes conséquences. Pour le dire de façon métaphorique, elles n'occupent pas la même place dans le réseau de nos représentations mentales ; en conséquence, elles n'invitent pas à faire les mêmes inférences.

C'est ce que semble suggérer Feynman à propos des différentes formulations de la loi de gravitation (loi de Newton, méthode du champ local et principe de minimum), dont il dit qu'elles sont « scientifiquement équivalentes », dans la mesure où « il est impossible de faire un choix, car il n'y a pas de manière de les départager expérimentalement puisque l'ensemble de leurs conséquences est identique » mais « psychologiquement différentes, et cela en deux sens » :

Premièrement, vous pouvez les aimer ou non, pour des raisons philosophiques ; à cela, l'exercice est le seul remède. Deuxièmement, elles sont différentes psychologiquement parce qu'elles ne sont pas du tout équivalentes quand il s'agit de découvrir de nouvelles lois.

Tant que la physique est incomplète, et que nous essayons de comprendre de nouvelles lois, les différentes formulations possibles peuvent fournir des indices de ce qui pourrait se passer dans d'autres circonstances. Elles ne sont alors plus équivalentes, psychologiquement, dans la mesure où elles nous suggèrent des idées sur l'aspect que pourraient revêtir les lois dans une situation plus générale. (Feynman, 1965, p. 53)

Naturellement, « préférer » la formulation newtonienne ou la formulation hamiltonienne ne signifie pas que l'on s'interdise l'utilisation d'une autre formulation : il ne s'agit pas de choisir l'une ou l'autre formulation à l'exclusion des autres, puisque chacune d'entre elles permet de déduire les autres. Préférer la formulation newtonienne, c'est plutôt considérer que la mécanique dans son ensemble est d'abord fondée sur des principes newtoniens.

En outre, considérer les différentes formulations de la mécanique comme l'expression de différentes versions ne doit pas nous faire oublier que chacune de ces formulations peut elle-même être comprise de diverses manières – faire l'objet de différentes versions. Les versions ne sont pas des ensembles de représentations exclusifs et isolés des autres représentations mentales d'un agent. On peut ainsi raisonnable-

ment affirmer que chaque expert a une certaine version de la mécanique – qui lui fait préférer, par exemple, une représentation en termes d'énergie à une représentation en termes de force –, dont dépend la manière dont il comprend et relie entre elles les différentes formulations.

Cette approche permet de rendre compte des différences entre les formulations comme d'authentiques différences conceptuelles, sans pour autant affirmer qu'elles sont contradictoires, incompatibles ou incommensurables : en tant que l'utilisation et la compréhension d'une théorie est toujours située, un agent, même expert, se représente toujours l'organisation des différents principes et leur lien au monde empirique sous une certaine perspective. Cela explique, comme je l'ai suggéré, à la fois la manière dont il utilise ces principes pour résoudre des problèmes et dont il les développe et les relie à d'autres domaines théoriques. Cela n'implique aucunement, cependant, de tracer une démarcation nette entre les formulations.

Bien au contraire, selon les différences auxquelles on accorde le plus d'importance – différences de langage mathématique, différences architectoniques, etc. –, on ne trace pas la même frontière entre les différentes formulations. En effet, alors que le principe de d'Alembert, en tant qu'il repose sur la description des systèmes au moyen de coordonnées généralisées et qu'il permet de déduire les équations de Lagrange, figure à bon droit dans un exposé de la mécanique analytique aux côtés du principe de Hamilton, il figure « du côté » de la mécanique newtonienne si l'on considère que cette dernière est caractérisée, plutôt que par son usage de coordonnées cartésiennes, par le fait qu'elle repose sur un principe différentiel exprimé au moyen du concept de force.⁸ En conséquence, selon ma version de la mécanique – et selon ma version de chacune de ses formulations –, je n'évalue pas leurs différences de la même manière. C'est là un des points essentiels que l'analyse des versions des sections 3 et 4 vise à montrer.

Pour finir, affirmer que les formulations newtonienne, lagrangienne et hamiltonienne d'aujourd'hui sont les expressions de différentes versions permet d'aborder sous un jour nouveau le problème historiographique que j'ai mentionné en introduction. En effet, parmi les versions qui ont fait l'objet de présentations explicites au cours de l'histoire, certaines, dont j'ai dit qu'elles étaient « innovantes » (voir page 98), ont donné lieu à la formulation de nouveaux principes, à l'introduction de nouveaux concepts, et à une réorganisation de l'architecture déductive de la théorie, qui s'accompagne souvent du développement des mathématiques nécessaires à l'établissement de ces nouveaux liens déductifs. Les formulations actuelles de la mécanique

⁸En effet, comme je l'ai souligné plusieurs fois, le principe de d'Alembert est lui-même un principe différentiel, et non intégral.

portent le nom des auteurs de certaines de ces versions particulièrement innovantes.

On peut considérer cela comme l'expression d'un jugement de la postérité selon lequel les formulations newtonienne, lagrangienne et hamiltonienne de la mécanique correspondent aux versions que Newton, Lagrange et Hamilton ont exprimées dans leurs présentations. Cela reviendrait à dire que les différences entre la présentation des *Principia* de Newton (1687) et la formulation newtonienne actuelle, ou entre la présentation de la *Mécanique analytique* de Lagrange (1788) et la formulation lagrangienne actuelle, sont des différences inessentiels au contenu de ces formulations. Autrement dit, cela signifierait que les innovations introduites par ces différents savants correspondent, diachroniquement, aux différences synchroniques entre les formulations d'aujourd'hui.

Évaluer ce jugement implique deux choses : d'abord, analyser la version que ces différents auteurs ont effectivement exprimée ; ensuite, identifier, au sein des formulations actuelles, ce que l'on considère essentiel à leur contenu. C'est un des buts de la suite de ce chapitre (sections 3 et 4) ; la prochaine section vise à préciser la méthode que j'entends y suivre.

2.4 Versions et histoire de la mécanique

La suite de ce chapitre est consacrée à l'examen de plusieurs versions de la mécanique, telles qu'elles ont été exprimées dans des présentations explicites. Cet examen va me permettre à la fois d'étayer l'hypothèse des versions au moyen d'exemples particuliers, et d'éclairer le problème historiographique de l'attribution des différentes formulations d'aujourd'hui à certains acteurs de l'histoire de la mécanique.

Rappelons ce problème. Les formulations standard présentées au chapitre 2 diffèrent en plusieurs points des travaux historiques des savants dont elles portent les noms. Prenons pour commencer l'exemple de la formulation newtonienne : elle est aujourd'hui exprimée en langage vectoriel et à l'aide d'équations différentielles, alors que toutes les démonstrations des *Principia* sont faites en langage géométrique et non à l'aide du calcul différentiel ; bien que les trois axiomes du mouvement que l'on a retenus comme les principes de la mécanique newtonienne figurent en tant que tels dans le texte de Newton, la deuxième loi n'y apparaît pas sous sa forme différentielle aujourd'hui standard, et la force et l'accélération n'y sont pas représentées par des quantités vectorielles (Newton, 1687, p. 11). Comme l'a montré Michel Blay (1992), c'est à Pierre Varignon (1700a,b) que l'on doit une réécriture des principes newtoniens à l'aide du formalisme différentiel, mis au point par Leibniz.⁹

⁹Pour une histoire de la mécanique newtonienne, voir (Dugas, 1950; Blay, 1992, 1995, 2002, 1993; Cohen, 1971, 1980; De Gandt, 1995; Wesfall, 1971; Whiteside, 1970).

C'est d'ailleurs cette rupture avec la mécanique *more geometrico* de Newton qui marque, selon Michel Blay, la naissance de la mécanique analytique.¹⁰ En effet, Michel Blay (1995, p. 108) affirme que la réécriture différentielle de Varignon « n'est pas qu'une affaire de langage », et qu'elle s'accompagne d'une « profonde refonte conceptuelle » de la théorie, par l'introduction des notions de vitesse dans chaque instant et de force accélératrice dans chaque instant.

Je ne développerai pas ici la question de la relation entre la présentation de Newton et la formulation newtonienne, ni celle de l'apport conceptuel de Varignon.¹¹ Je souhaite simplement, par cet exemple, suggérer le lien, que le reste de ce chapitre vise à étudier en détail, entre le jugement historique porté sur la contribution de tel ou tel acteur du développement de la mécanique et l'adoption d'une certaine version de cette théorie : ce jugement porté par Michel Blay repose sur une certaine définition de la mécanique analytique, à la lumière de laquelle il situe le moment où une authentique nouveauté conceptuelle est introduite. Autrement dit, c'est la manière dont l'historien Blay comprend la mécanique – c'est sa version de la mécanique – qui le conduit à juger que l'apport de Varignon est conceptuel et non une simple réécriture formelle des principes newtoniens. Ainsi, en raison même du fait que le développement historique de la mécanique ne consiste pas en des changements logico-empiriques, et est, en ce sens, purement formel, l'historien jugera que la nouveauté introduite par tel ou tel acteur de son histoire est authentiquement conceptuelle ou non en fonction de sa propre version de la mécanique et de la manière dont il évalue la différence entre ses formulations.

C'est à un examen approfondi de la manière dont ce problème se pose pour le cas de Lagrange – et non de Newton – que la suite de ce chapitre est consacrée. Je ne chercherai pas tant à répondre à la question de l'apport authentique de Lagrange à la mécanique et à la légitimité de l'attribution de son nom à l'une des formulations analytiques d'aujourd'hui qu'à mettre en lumière, par un examen de la manière dont certains y ont effectivement répondu – implicitement ou explicitement –, la difficulté de trancher nettement entre développement formel et avancée conceptuelle.

J'ai choisi de me concentrer sur le cas de Lagrange et de la formulation lagran-

¹⁰Pour une histoire du calcul infinitésimal et de la naissance de la mécanique analytique, voir (Blay, 1992, 2002).

¹¹Signalons cependant que, malgré cette différence de forme mathématique (démonstrations géométriques *versus* équations différentielles), on peut, contre la thèse de Michel Blay, affirmer que cette nouveauté mathématique n'est pas le signe d'une version très innovante, dans la mesure où les démonstrations géométriques de Newton avaient probablement une valeur pédagogique, et où on peut lui attribuer une représentation en termes de calcul différentiel, dont il se dispute la paternité avec Leibniz.

gienne pour plusieurs raisons. En premier lieu, notons que ce choix n'est pas véritablement exclusif : étudier la version qu'un individu a de la mécanique lagrangienne, c'est étudier, par cette « fenêtre », la manière dont il se représente l'ensemble de la mécanique et l'articulation de ses différents principes.

Par ailleurs, contrairement au cas de la mécanique newtonienne, les différences de langage mathématique entre la présentation de 1788 et la formulation lagrangienne actuelle sont négligeables. En conséquence, le problème est assez différent. D'une part, les différences entre la présentation de Lagrange et la formulation lagrangienne semblent moindres que dans le cas de la formulation newtonienne, puisque les équations lagrangiennes du mouvement ont à peu de choses près la forme que leur donnait Lagrange ; elles sont, comme on le verra, plus difficiles à caractériser d'emblée comme des différences « formelles ». D'autre part, le problème de la réécriture des principes de Newton par Varignon – et celui, corrélatif du « newtonianisme » de Newton – est un problème d'histoire conceptuelle qui ne concerne pas la manière dont on comprend aujourd'hui la mécanique et ses différentes formulations, dans la mesure où la présentation géométrique de Newton n'est plus utilisée depuis la réécriture varignonienne. En revanche, les différentes manières de comprendre et de présenter la formulation lagrangienne posent non seulement une question d'histoire conceptuelle, mais aussi des questions relatives à la compréhension que les différents individus, y compris après Lagrange, ont de la mécanique analytique.

On considère, et c'est la raison principale de mon choix, que Lagrange est le père de ce que l'on appelle « mécanique analytique » – le titre même de son ouvrage de 1788 semble autoriser ce jugement. Cependant, la définition de l'expression « mécanique analytique » est elle-même assez variable. Aujourd'hui – mais il n'en a pas toujours été ainsi –, on considère souvent que les traits distinctifs les plus remarquables de la mécanique analytique sont de reposer sur un principe variationnel et d'être une mécanique de l'énergie, et non de la force. Or, comme l'analyse de la version de Lagrange le montrera, cette assimilation entre mécanique analytique (lagrangienne), mécanique variationnelle et mécanique énergétique est loin d'aller de soi.

Dans la section 3, j'examine la version de Lagrange, telle qu'elle ressort de la *Mécanique analytique* de 1788. Le statut qu'il accorde aux différents principes, ainsi que le jugement explicite qu'il porte sur sa propre contribution à l'histoire de la mécanique, me conduiront à remettre en question la légitimité de l'attribution de la formulation lagrangienne actuelle à Lagrange. Cette conclusion, comme on le verra, est d'ailleurs conforme au jugement de plusieurs historiens et de certains des successeurs de Lagrange (comme Hamilton), qui considèrent que Lagrange a essentiellement procédé à un développement formel de la mécanique, et non à une refonte conceptuelle.

L'examen, à la section 4, de plusieurs présentations proposées par des savants et

des philosophes postérieurs à Lagrange et à Hamilton – en particulier au cours du dix-neuvième siècle – me permettra de nuancer cette dernière conclusion. On verra en effet que, selon leur version de la mécanique et, plus précisément, selon le statut et la signification qu'ils accordent à la mécanique analytique (lagrangienne), ils n'évaluent pas de la même manière les différences entre formulations et ne tracent pas la même frontière entre elles ; cela débouche, comme on le verra, sur des jugements différents à propos de la contribution historique de Lagrange.

3 La version de Lagrange (1788)

Lagrange est généralement considéré comme le père de la mécanique analytique. Comme on l'a vu, cette dernière se distingue de la mécanique newtonienne sous plusieurs aspects. Les équations de la mécanique analytique, sous leur forme lagrangienne ou hamiltonienne, représentent l'état des systèmes mécaniques au moyen de coordonnées généralisées ; on peut les obtenir à partir d'un principe variationnel – le principe de Hamilton – et elles gouvernent la dynamique des systèmes au moyen d'une quantité scalaire qui a la dimension d'une énergie.

Plusieurs raisons justifient que l'on attribue à Lagrange la paternité de la mécanique analytique ainsi conçue. Celui-ci, à la suite d'Euler, a en effet apporté une contribution majeure au développement du calcul variationnel.¹² Il est notamment connu pour avoir donné la première formulation mathématique rigoureuse et dépourvue de toute signification métaphysico-théologique du principe de moindre action¹³, suffisamment générale pour n'être pas restreinte, comme celle d'Euler, aux systèmes

¹²Pour une histoire du calcul variationnel, voir (Dahan-Dalmedico, 1988, 1990; Fraser, 1983, 1985b, 1994, 1997; Barroso Filho et Comte, 1988; Panza, 1990, 2003; Galletto, 1991; Barroso Filho, 1994; Martin-Robine, 2006; Goldstine, 1980).

¹³Pour une histoire du principe de moindre action, voir (Pulte, 1989; Martin-Robine, 2006). Ce principe était initialement destiné à offrir un fondement métaphysique ou théologique à l'ensemble de la physique, permettant de fournir des explications par les causes finales et non par les seules causes efficientes. La première formulation de la mécanique au moyen d'un tel principe a été proposée par de Maupertuis (1746). L'action y est alors définie comme le produit de la masse, de la vitesse et de la distance (*mvs*). Le manque de rigueur de cette définition a été souligné maintes fois, et personne ne prétend aujourd'hui que le principe de Maupertuis est authentiquement scientifique. Cependant, comme le note Anouk Barberousse (2008), il permet malgré tout d'introduire un nouveau concept, celui d'action, dans la mécanique, et d'ouvrir la voie à une formulation de la mécanique qui se passerait de la notion de force, jugée douteuse. En outre, comme le montre Maël Pegny (2005), à la suite de Helmut Pulte (1989), l'influence de Maupertuis sur les travaux d'Euler est loin d'être négligeable ; l'opposition classique entre un Maupertuis dont les seules motivations sont théologiques et un Euler appliquant une méthode scientifique rigoureusement inductive est à nuancer. À ce sujet, voir (Pegny, 2005), qui examine en détail les motivations d'Euler.

conservatifs.¹⁴ En distinguant une forme étroite – restreinte aux systèmes conservatifs – et une forme générale du principe, Lagrange permet non seulement de déduire de ce principe des solutions à des problèmes particuliers (en déduisant des fonctions) mais également un système d'équations différentielles équivalentes aux équations newtoniennes (connues aujourd'hui sous le nom d'« équations d'Euler-Lagrange »). Ces équations préfigurent les équations aujourd'hui appelées « lagrangiennes » qui sont au centre de la formulation lagrangienne actuelle, dans le cadre de laquelle elles sont déduites d'un principe variationnel, le principe de Hamilton, et formulées, comme lui, au moyen d'une fonction scalaire L (appelée « lagrangien ») qui a la dimension d'une énergie.

Pourtant, la version de la mécanique exprimée par Lagrange (1788) est assez différente de celle que suggèrent la formulation lagrangienne et l'opinion générale qui la range dans la classe des formulations variationnelles – c'est-à-dire des formulations de la mécanique qui sont fondées sur un principe variationnel. Un examen de l'architecture logique de la *Mécanique analytique*, ainsi que des jugements explicites portés par Lagrange sur le statut des différents principes et sur sa propre contribution au développement de la mécanique révèle en effet un Lagrange bien plus newtonien – ou, plus précisément, d'alembertien – et attaché à une mécanique de la force que l'association de son nom à celui de la mécanique variationnelle ne le laisse imaginer. Autrement dit, comme on va le voir, si Lagrange est bien le père de la mécanique *analytique*, l'assimilation de cette dernière à une mécanique *variationnelle* dont le concept central est l'énergie ne va pas de soi.¹⁵

3.1 L'architecture logique de la *Mécanique analytique* (1788)

Avant d'examiner la place que le principe de moindre action occupe dans la *Mécanique analytique*, il convient de dire un mot des travaux de Lagrange sur le calcul variationnel. Toute une partie de son œuvre vise en effet à développer ce qu'Amy Dahan-Dalmedico (1988, 1990) appelle un « style de mathématisation », en l'occurrence le formalisme variationnel, dans lequel il voit le moyen de résoudre tout un ensemble de problèmes physiques.

La fin de la décennie 1750 est ainsi consacrée au développement de la méthode variationnelle et à son application à différents problèmes physiques.¹⁶ À cette époque,

¹⁴Pour la formulation du principe de moindre action par Euler (1744) et pour l'influence de ce dernier sur les travaux de Lagrange, voir (Pulte, 1989; Pegny, 2005).

¹⁵Mis à part les quelques passages originaux de Lagrange que je cite, le matériau historique de cette analyse est presque entièrement emprunté aux sources secondaires que je mentionnerai au cours de l'exposé.

¹⁶L'année 1760 voit ainsi paraître deux œuvres de toute première importance pour le développe-

le très jeune Lagrange place un grand espoir dans le principe de moindre action, duquel il montre que l'on peut déduire un ensemble d'équations différentielles équivalentes aux équations newtoniennes. Comme en témoigne ce passage d'une lettre à Euler du 19 mai 1756, il envisage initialement d'en faire le principe fondamental de la mécanique :

Au sujet du principe de moindre quantité d'action, je pense pour ma part qu'on peut le considérer comme la clé universelle de tous les problèmes, tant statiques que dynamiques, pour les questions relevant soit du mouvement des corps – quel que soit leur nombre et quelle que soit la manière dont ils sont liés entre eux –, soit de l'équilibre et du mouvement des fluides quelconques : le fait me paraît évident, si l'on rajoute aux remarquables résultats concernant son application à la mécanique, que vous avez déjà publiés en maints endroits, ceux assez peu nombreux, que je vous ai en partie communiqués, gardant le reste par devers moi. (Cité dans Pegny, 2005, p. 69)

Pourtant, ce jugement enthousiaste du jeune mathématicien va progressivement être contrebalancé par une appréciation assez différente du principe de moindre action et de son statut dans l'architecture globale de la mécanique : dès les *Recherches sur la libration de la Lune* (Lagrange, 1764), ce principe est présenté comme une simple conséquence du principe de d'Alembert et cesse d'occuper la place centrale qui lui avait initialement été attribuée. Finalement, un regard sur l'architecture logique de la *Mécanique analytique* (voir ci-dessous, figure 8 page 121, la table des matières de l'ouvrage), ainsi que sur les jugements explicites que Lagrange y porte sur l'ordre et le statut des différents principes, ébranle fortement l'idée selon laquelle il aurait conçu la mécanique comme fondée sur un principe variationnel. Les deux parties de l'ouvrage, statique et dynamique, sont en effet fondées sur des principes différentiels, et non sur des principes variationnels.

La première moitié de l'ouvrage de 1788 est consacrée à la statique, c'est-à-dire aux principes qui gouvernent l'équilibre des systèmes mécaniques. Lagrange en expose les trois principes fondamentaux, à savoir ceux de l'équilibre du levier, de la composition des mouvements, et enfin des vitesses virtuelles.¹⁷ C'est ce dernier qu'il érige

ment du calcul variationnel et son application aux problèmes de mécanique : l'*Essai sur une nouvelle méthode pour déterminer les maximas et les minims des formules intégrales indéfinies* (Lagrange, 1760b) et l'*Application de la méthode proposée dans le mémoire précédent à la solution de différents problèmes de mécanique* (Lagrange, 1760a).

¹⁷Il énonce le principe des vitesses virtuelles comme suit : « Si un système quelconque de tant de corps ou points que l'on veut tirés, chacun par des puissances quelconques, est en équilibre, et qu'on donne à ce système un petit mouvement quelconque, en vertu duquel chaque point parcourt un espace infiniment petit qui exprimera sa vitesse virtuelle ; la somme des puissances, multipliées chacune par l'espace que le point où elle est appliquée, parcourt suivant la direction de cette même

au rang de principe fondamental, qu'on peut « regarder comme une espèce d'axiome de mécanique » (Lagrange, 1788, p. 12), insistant sur sa *simplicité* et sa *généralité* :

Soit qu'on regarde le Principe des vitesses virtuelles comme une propriété générale de l'équilibre, ainsi que l'a fait Galilée ; soit qu'on veuille le prendre avec Descartes et Wallis pour la vraie cause de l'équilibre, il faut avouer qu'il a toute la simplicité qu'on peut désirer dans un principe fondamental ; et nous verrons plus bas combien ce principe est recommandable par sa généralité. [...]

Et en général je crois pouvoir avancer que tous les principes généraux qu'on pourrait peut-être encore découvrir dans la science de l'équilibre, ne seront que le même principe des vitesses virtuelles, envisagé différemment, et dont ils ne différeront que dans l'expression.

Au reste, ce principe est non seulement en lui-même très simple et très général ; il a de plus l'avantage précieux et unique de pouvoir se traduire en une formule générale qui renferme tous les problèmes qu'on peut proposer sur l'équilibre des corps. Nous allons exposer cette formule dans toute son étendue ; nous tâcherons même de la présenter d'une manière encore plus générale qu'on ne l'a fait jusqu'à présent, et d'en donner des applications nouvelles. (Lagrange, 1788, pp. 10-12)

Dans la seconde section de la statique, intitulée « Formule générale pour l'équilibre d'un système quelconque de forces ; avec la manière de faire usage de cette formule », il montre comment « réduire ce principe en formule » (Lagrange, 1788, p. 12), c'est-à-dire comment transformer l'énoncé du principe en une équation permettant de résoudre les problèmes de statique.

En quoi consiste l'innovation de Lagrange ici ? Il invente une méthode permettant, à partir d'un principe connu, auquel il ne semble ajouter aucune signification nouvelle, d'obtenir une formule générale de laquelle on peut déduire les équations qui gouvernent la statique des systèmes mécaniques. Cette méthode, en termes modernes, consiste à représenter l'état d'un système à l'aide de coordonnées généralisées, et donc à exprimer le principe des vitesses virtuelles en fonction des contraintes du système. Lagrange n'ajoute donc aucun *principe* nouveau, mais propose une nouvelle *méthode* mathématique plus efficace.

La seconde moitié de l'ouvrage est consacrée à la dynamique, dont Lagrange considère qu'elle doit être fondée sur la statique. Le principe premier de la dynamique, dans ce cadre, est donc l'extension du principe des vitesses virtuelles aux cas dynamiques, c'est-à-dire le principe de d'Alembert.

Le traité de Dynamique de M. d'Alembert qui parut en 1743, mit fin à ces

puissance, sera toujours égale à zéro, en regardant comme positifs les petits espaces parcourus dans le sens des puissances, et comme négatifs les espaces parcourus dans un sens opposé. » (Lagrange, 1788, pp. 10-11, italiques originales)

espèces de défis¹⁸, en offrant une méthode directe et générale pour résoudre, ou du moins pour mettre en équations tous les problèmes de dynamique que l'on peut imaginer. Cette méthode réduit toutes les lois du mouvement des corps à celles de leur équilibre, et ramène ainsi la dynamique à la statique. (Lagrange, 1788, p. 179)

Cependant, tel quel,

[c]e principe ne fournit pas immédiatement les équations nécessaires pour la solution des différents problèmes de dynamique, mais il apprend à les déduire des conditions de l'équilibre. Ainsi en combinant ce principe avec les principes ordinaires de l'équilibre du levier, ou de la composition des forces, on peut toujours trouver les équations de chaque problème à l'aide de quelques constructions plus ou moins compliquées. (Lagrange, 1788, p. 180)

Tout comme il avait traduit le principe des vitesses virtuelles en une formule analytique permettant de résoudre les problèmes de statique, en combinant le principe des vitesses virtuelles exposé dans la statique avec le principe de d'Alembert, Lagrange obtient « une méthode semblable pour les problèmes de dynamique, et qui [a] les mêmes avantages » (Lagrange, 1788, p. 181).

De cette manière on pourra exprimer analytiquement les termes dont il s'agit, et l'on aura une formule générale pour le mouvement des corps, laquelle renfermera la solution de tous les problèmes de dynamique, et dont le simple développement donnera les équations nécessaires pour chaque problème, [...]. (Lagrange, 1788, p. 182)

Le principe de d'Alembert, une fois exprimé en fonction des contraintes du système, permet bien, comme on l'a vu au chapitre 2, d'obtenir les équations du mouvement.

Jusqu'ici, pas de trace d'un principe de minimum, ni d'un concept énergétique.¹⁹

¹⁸Les défis en question sont les problèmes de dynamique auxquels se sont attaqués Euler, les Bernoulli, Clairaut, et auxquels ils ont répondu par des versions successives du principe qu'« il était réservé à M. d'Alembert d'envisager [...] d'une manière générale, [en lui donnant] toute la simplicité et la fécondité dont il pouvait être susceptible. ». Il énonce ce principe comme suit : « Si plusieurs corps tendent à se mouvoir avec des vitesses et des directions, qu'ils soient forcés de changer à cause de leur action mutuelle, on peut regarder ces mouvements comme composés de ceux que les corps prendront réellement, et d'autres mouvements qui sont détruits ; d'où il suit que ces derniers doivent être tels que les corps animés de ces seuls mouvements se fassent équilibre. » (Lagrange, 1788, p. 180) Pour une analyse de la formulation originale du principe par d'Alembert, voir (Fraser, 1985a; Darrigol, 2007).

¹⁹Le concept d'énergie n'existe pas en tant que tel à l'époque de Lagrange. Cependant, comme le note Maël Pegny (2005), on distingue clairement de la force la *vis viva* et la *vis potentialis*, ancêtres respectifs de l'énergie cinétique et de l'énergie potentielle. Lagrange, donc, « aurait pu » leur donner un rôle central. C'est précisément le fait qu'il soit formulé au moyen de concepts énergétiques qui va permettre au principe de moindre action de devenir central dans d'autres domaines que celui de la seule mécanique, auquel la notion de force est limitée. Pour une histoire des concepts physiques

Le principe de d'Alembert, institué en principe fondamental de la dynamique, peut être décrit comme une reformulation ingénieuse de la deuxième loi de Newton qui, jointe à la méthode analytique mise au point par le mathématicien Lagrange, offre un outil unique permettant d'obtenir toutes les équations du mouvement pour un système mécanique, même sujet à de nombreuses contraintes. Ce progrès considérable dans la simplicité et la généralité, comme le souligne à plusieurs reprises Lagrange, ne semble donc pas modifier la signification physique de la mécanique : aucun principe, ni aucun concept authentiquement nouveau n'est ajouté à l'édifice, qui a simplement gagné en élégance et rend possible la résolution de problèmes mécaniques plus nombreux.

Quelle est donc la place du principe de moindre action dans l'architectonique de la mécanique de Lagrange en 1788 ? Avant de déduire les équations du mouvement du principe de d'Alembert reformulé en termes de coordonnées généralisées, Lagrange examine plusieurs conséquences de ce principe. Parmi elles, au même titre que les principes de conservation, figure le « principe de la moindre quantité d'action » :

Mais un des plus grands avantages de cette formule, est d'offrir immédiatement les équations générales qui renferment les principes, ou théorèmes connus sous les noms de *conservation des forces vives*, de *conservation du mouvement du centre de gravité*, de *conservation du moment du mouvement de rotation*, ou *principe des aires*, et de *principe de la moindre quantité d'action*. Ces principes doivent être regardés plutôt comme des *résultats généraux*²⁰ des lois de la dynamique, que comme des *principes primitifs*²¹ de cette science, mais *étant souvent employés comme tels*²² dans la solution des problèmes, nous croyons devoir en dire aussi un mot, en indiquant en quoi ils consistent, et à quels auteurs ils sont dus, pour ne rien laisser à désirer dans cette exposition préliminaire des principes de la dynamique. (Lagrange, 1788, p. 182)

Tout en insistant sur sa propre contribution à l'histoire de ce principe²³, Lagrange ajoute :

au XIX^e siècle, voir (Harman, 1982a; Darrigol, 2001).

²⁰ Je souligne.

²¹ Je souligne.

²² Je souligne.

²³ Lagrange énonce ainsi le principe de moindre action qu'il attribue à Euler (1744) : « dans les trajectoires décrites par des forces centrales, l'intégrale de la vitesse multipliée par l'élément de la courbe, fait toujours un *maximum* ou un *minimum*. [...] Cette propriété que M. Euler n'avait reconnue que dans le mouvement des corps isolés, je l'ai étendue depuis au mouvement des corps qui agissent les uns sur les autres d'une manière quelconque, et il en a résulté ce nouveau principe général, que la somme des produits des masses par les intégrales des vitesses multipliées par les éléments des espaces parcourus, est constamment un *maximum* ou un *minimum*. » (Lagrange, 1788, pp. 188-189)

Tel est le principe auquel je donne ici, quoique improprement, le nom de *moindre action*, et que je regarde non comme un principe métaphysique, mais comme un résultat simple et général des lois de la mécanique. On peut voir dans le tome II des Mémoires de Turin, l'usage que j'en ai fait pour résoudre plusieurs problèmes difficiles de dynamique. Ce principe combiné avec celui de la conservation des forces vives, et développé suivant les règles du calcul des variations, donne directement toutes les équations nécessaires pour la solution de chaque problème; et de là naît une méthode également simple et générale pour traiter les questions qui concernent le mouvement des corps; mais *cette méthode n'est elle-même qu'un corollaire de celle qui fait l'objet de la seconde partie de cet ouvrage, et qui a en même temps l'avantage d'être tirée des premiers principes de la mécanique*²⁴. (Lagrange, 1788, p. 189)

Lagrange reconnaît donc bien la possibilité de fonder la dynamique sur le principe variationnel auquel il a lui-même donné une forme générale nouvelle, rendant possible la déduction des équations du mouvement. Son jugement, cependant, est sans équivoque : ce principe n'est pas un principe primitif de la théorie, c'est une conséquence intéressante de la formule générale obtenue à partir du principe de d'Alembert, seul principe à mériter véritablement de figurer au fondement de l'édifice de la dynamique.

Cette préférence de Lagrange pour le principe de d'Alembert sur le principe de moindre action peut s'expliquer de plusieurs manières. La première raison en est que le principe de moindre action, tel qu'il est formulé par Lagrange, est moins général que celui de d'Alembert. Il ne permet pas de traiter les systèmes non holonomes (en termes modernes²⁵), les forces non intégrables, et les chocs durs.²⁶

Une fois que Hamilton aura levé cette restriction du principe en définissant l'action d'un système en fonction des points initial et final de la trajectoire et de l'énergie du système, le principe de moindre action offrira une méthode encore plus « simple et générale », pour reprendre les termes de Lagrange, que celle fondée sur le principe de d'Alembert. Tant que ce n'est pas le cas, il n'y a aucune raison de le préférer à cette dernière méthode : l'édifice de la mécanique n'en gagnerait pas en élégance ni en simplicité. En présentant ainsi la mécanique, on peut déduire tout ce qu'elle dit à partir d'un seul et unique principe.

Cette raison est, à elle seule, suffisante à Lagrange pour ne pas chercher à fonder la mécanique sur un principe nouveau. Un de ses buts est en effet, comme le souligne Amy Dahan-Dalmedico (1990), de donner « une synthèse vaste et générale, permettant d'obtenir a priori des méthodes, démonstrations et procédures uniformes ». Je propose cependant, dans la section suivante, d'approfondir l'analyse des raisons qui

²⁴ Je souligne.

²⁵ Voir l'annexe A.5.

²⁶ Merci à Olivier Darrigol pour ces remarques. Voir aussi (Darrigol, 2001).

poussent Lagrange à ne pas faire de sa mécanique analytique une mécanique variationnelle proprement dite.

3.2 Version de la mécanique et « idéal scientifique » de Lagrange

On peut considérer que l'architecture logique que Lagrange donne à la mécanique – et, plus généralement, sa version de la mécanique – est le résultat de ce qu'Amy Dahan-Dalmedico (1990) appelle son « idéal scientifique », et qu'elle caractérise comme suit :

- Symétrie dans les calculs
- Appel privilégié à l'algèbre et sa marche régulière et uniforme
- Méthodes fixes et générales
- Formalisme des procédures
- Expulsion de toute trace de « métaphysique » (au sens courant des géomètres du XVIII^e) dans les théories, etc. (Dahan-Dalmedico, 1990, p. 98)

Les quatre premiers points ont été illustrés par ce qui précède. Le cinquième invite à approfondir l'analyse : non seulement le but principal poursuivi par Lagrange est de fournir une méthode générale et, de ce fait, suffit à choisir la méthode partant du principe de d'Alembert, mais encore sa volonté d'« expuls[er] [...] toute trace de métaphysique » lui évite d'être séduit, comme l'avaient été Euler²⁷, et, avant lui, Maupertuis, par la refondation de la mécanique sur un principe susceptible d'appeler une interprétation téléologique. En en faisant une simple conséquence du principe de d'Alembert, et non un principe fondamental, Lagrange contribue à débarrasser le principe de moindre action de toute empreinte métaphysique : il n'apporte en effet aucune signification nouvelle au système.

On peut suggérer que Lagrange a donc « tout à gagner » à maintenir le principe de d'Alembert au fondement de sa mécanique : cela lui permet d'avoir un outil plus général, qui ne requiert pas que les forces soient intégrables, et cela lui évite un énoncé trop chargé de métaphysique.²⁸ Autrement dit : ce n'est peut-être pas seulement parce que le principe de moindre action sous sa forme lagrangienne est moins général que le principe de d'Alembert que Lagrange le « regarde non comme un principe métaphysique, mais comme un résultat simple et général des lois de la mécanique », mais aussi *parce qu'il le regarde ainsi* qu'il n'a aucune raison de chercher à le placer au fondement de sa mécanique et d'opérer ainsi une révolution conceptuelle : d'autres que lui, comme Maupertuis, avaient sacrifié à l'idéal de rigueur et de

²⁷Voir (Pegny, 2005) pour une analyse des motivations de Euler ; il montre notamment qu'Euler cherchait dans le principe de moindre action la clé d'une explication finale des phénomènes, permettant de réorganiser le champ de la connaissance.

²⁸Voir (Dahan-Dalmedico, 1990).

simplicité mathématique pour tenter de fonder la physique entière sur un principe de minimum.

En outre, au delà même de sa méfiance vis-à-vis d'une possible interprétation métaphysique de ce qui n'est dans son système qu'une formule mathématique déduite d'un principe plus fondamental, on peut affirmer que, de son propre aveu, Lagrange cherche assez peu à explorer la signification physique des principes de la mécanique.

À en croire les jugements explicites que Lagrange porte sur sa propre contribution au développement de la mécanique, celle-ci est essentiellement formelle : elle contribue à organiser les principes et à les reformuler d'une manière commode et féconde, mais son travail ne prétend pas être une enquête systématique sur les conséquences des lois de la mécanique, c'est-à-dire sur leur signification physique ; il consiste plutôt, selon lui, dans le calcul d'implications purement mathématiques. C'est ce dont témoigne clairement le célèbre « Avertissement » que Lagrange place en tête de sa *Mécanique analytique*, que Hamilton qualifiera de « poème scientifique par le Shakespeare des mathématiques » :

On a déjà plusieurs Traités de Mécanique, mais le plan de celui-ci est entièrement neuf. Je me suis proposé de réduire la théorie de cette science, et l'art de résoudre les problèmes qui s'y rapportent, à des formules générales, dont le simple développement donne toutes les équations nécessaires pour la solution de chaque problème. [...]

On ne trouvera point de figures dans cet ouvrage. Les méthodes que j'y expose ne demandent ni constructions, ni raisonnements géométriques ou mécaniques, mais seulement des opérations algébriques, assujetties à une marche régulière et uniforme. Ceux qui aiment l'analyse, verront avec plaisir la mécanique en devenir une nouvelle branche, et me sauront gré d'en avoir étendu ainsi le domaine.
(Lagrange, 1788, p. v-vi)

À l'en croire, l'œuvre de Lagrange est donc en quelque sorte celle d'un mathématicien. Par ces propos, il illustre en effet assez bien les considérations de Feynman sur les mathématiciens, qui, selon lui, « s'intéressent seulement à la structure du raisonnement, et ne se préoccupent pas vraiment de ce dont ils parlent », alors que « le physicien donne un sens à chacune de ses phrases » (Feynman, 1965, p. 55). D'après l'« Avertissement » de la *Mécanique analytique*, c'est l'analyse qui s'y trouve enrichie, et non la mécanique.²⁹ Cette dernière a simplement été dotée d'outils puissants, mais

²⁹ Je ne prétends pas ici affirmer que Lagrange se comporte uniquement en mathématicien : la critique que je propose, à la fin de ce chapitre, de la distinction entre un développement formel et une avancée conceptuelle montrera au contraire que, dans certains cas, des développements « purement formels » comptent comme des avancées conceptuelles. Dans ce cas précis, les développements mathématiques de Lagrange sont le plus souvent destinés à rendre la mécanique applicable à de nouvelles classes de problèmes : théorie des perturbations planétaires, stabilités, systèmes vibratoires,

sa signification physique en est inchangée.

En conséquence, sa signification physique lui est tout entière fournie par les principes newtoniens, dont le principe des travaux virtuels et le principe de d'Alembert sont présentés comme des réécritures commodes. On peut donc affirmer que la mécanique lagrangienne est une mécanique de la force, puisque c'est le concept central de ces principes : ils représentent et expliquent les phénomènes du mouvement en décrivant les forces qui en sont la cause.

Or, et c'est là un deuxième argument en faveur de l'affirmation selon laquelle Lagrange incarne bien la figure du mathématicien tel que Feynman le dépeint, il ne consacre pas le moindre examen critique au concept de force, pourtant tenu en suspicion par nombre de ses contemporains.³⁰ La définition qu'il en donne dès la première page de l'ouvrage, manifestant par là qu'il s'agit bien du concept premier de la mécanique, est étonnamment sommaire et peu satisfaisante pour quiconque y cherche une signification claire :

On entend en général par *force* ou *puissance* la cause, quelle qu'elle soit, qui imprime ou tend à imprimer du mouvement au corps auquel on la suppose appliquée; et c'est aussi par la quantité du mouvement imprimé, ou prêt à imprimer, que la force ou puissance doit s'estimer. Dans l'état d'équilibre la force n'a pas d'exercice actuel; elle ne produit qu'une simple tendance au mouvement; mais on doit toujours la mesurer par l'effet qu'elle produirait si elle n'était pas arrêtée. En prenant une force quelconque, ou son effet pour l'unité, l'expression de toute autre force n'est plus qu'un rapport, une quantité mathématique qui peut être représentée par des nombres ou des lignes; c'est sous ce point de vue que l'on doit considérer les forces dans la mécanique. (Lagrange, 1788, pp. 1-2)

Lagrange présente donc sa mécanique comme une réécriture et une confirmation des principes newtoniens, et adopte sans autre examen le concept de force comme concept premier. Sur ce point, il n'a rien à ajouter à ce qu'a dit Newton : il peut mettre en œuvre son idéal de simplicité et de généralité, sans interroger la signification physique des principes.

Ce constat s'accorde d'ailleurs parfaitement avec le jugement que Lagrange porte sur l'histoire des sciences :

Dès mes premières études j'avais conçu pour d'Alembert une admiration passionnée, et je l'ai toujours conservée, parce que c'est lui qui, dans le siècle dernier, a fait le plus de brillantes découvertes. Je conviens cependant qu'on étudiera plutôt Euler dans tous les temps, et avec raison, parce qu'il a mieux écrit. Ce

acoustiques, etc. (Merci à Olivier Darrigol pour ces précisions) Ce que je souligne ici est que, en affirmant qu'il n'a pas modifié la signification physique de la mécanique, Lagrange évalue lui-même sa contribution à la mécanique comme un développement formel.

³⁰ Voir (Pulte, 1989, pp. 237-238).

sont là mes deux grands hommes, ceux dont je fais le plus de cas après Newton ;
 mais tout le monde ne peut pas être aussi *heureux* que Newton... (cité dans
 Pulte, 1989, p. 232, note 154)

Newton est le plus heureux des hommes et personne ne peut le dépasser, dans la mesure où, selon Lagrange, *tout est dit* dans les *Principia* : les travaux ultérieurs – de d'Alembert, d'Euler, et de Lagrange lui-même – en sont des réécritures formelles, qui les rendent applicables à des domaines pour lesquels ils étaient pratiquement inutilisables, mais qui n'en modifient pas la signification physique.

Pour résumer, la mécanique selon Lagrange – la version de Lagrange – est fondée sur le principe de d'Alembert, et tire sa signification du concept de force. Le principe de moindre action y figure, à titre de théorème. De cela, je voudrais tirer deux ensembles de conclusions.

En premier lieu, je voudrais insister sur la chose suivante : la version de Lagrange, qui se manifeste dans le statut qu'il attribue aux différents principes et dans la signification qu'il leur donne, est profondément liée à ce qu'Amy Dahan appelle son « idéal scientifique », et plus particulièrement son exigence formelle (sa recherche de la « simplicité » et de la « généralité »), ainsi qu'à son acceptation sans examen d'une représentation des phénomènes du mouvement en termes de force, et enfin à son jugement sur l'histoire des sciences et sur sa propre contribution.

Mon but n'est pas ici de reconstituer l'histoire intellectuelle et la psychologie de Lagrange, à laquelle notre accès est fort limité. Je ne me prononcerai pas sur la question de savoir, par exemple, si c'est son admiration pour Newton et un engagement initial en faveur du concept de force qui le conduisent à se limiter à une réécriture formelle des principes newtoniens, ou, inversement, si c'est son esprit mathématique et anti-métaphysique qui le détournent des questions relatives à la signification physique des principes, le faisant accepter sans examen le concept de force. Ce qui m'intéresse ici n'est pas de répondre à ce genre de questions, mais d'insister sur le fait que, quelles que soient ici la cause et la conséquence, la conviction de Lagrange qu'il n'y a rien à ajouter à la signification physique de la mécanique newtonienne et son jugement sur sa propre contribution – selon lequel elle est strictement formelle – sont indissociables.

En second lieu, l'examen de la version de Lagrange semble remettre en cause l'idée selon laquelle il serait le père de la formulation lagrangienne d'aujourd'hui, en tant qu'elle est une branche de la mécanique variationnelle. Quelle que soit l'importance de sa contribution – et bien qu'il ait effectivement développé le calcul variationnel et montré que l'on pouvait déduire les équations du mouvement d'un principe de minimum –, sa mécanique n'était pas une mécanique *variationnelle*, au sens où elle n'était pas fondée sur un principe variationnel ; le concept d'énergie, ou de force vive, n'y joue pas de rôle fondamental. Comme on va le voir, le saut conceptuel qui consiste

à réorganiser entièrement l'architecture de l'édifice et à lui permettre de devenir une mécanique de l'énergie a été opéré par Hamilton (1834). Pour clore cette analyse de la version de Lagrange, je me propose, dans la section suivante, d'examiner les conséquences de ce constat pour l'appréciation rétrospective de la contribution de Lagrange.

3.3 Lagrange (1788) et Hamilton (1834) : développement formel contre avancée conceptuelle ?

Comme on l'a vu, une des raisons du maintien du principe de d'Alembert au fondement de la mécanique est que le principe de moindre action, jusqu'à sa reformulation par Hamilton, est moins général que le principe de d'Alembert. Le principe variationnel qui est au fondement de la formulation lagrangienne actuelle est en effet le principe de Hamilton : le savant irlandais, en définissant l'action en fonction des points initial et final de la trajectoire et de l'énergie³¹ du système, lève la restriction qui pesait sur le principe de moindre action dans sa version lagrangienne.³²

Comme je l'ai suggéré à la section 1 du présent chapitre, Hamilton, en établissant de nouveaux liens déductifs au sein de la mécanique et en modifiant considérablement l'architecture logique de la mécanique par une refondation sur le principe de moindre action, donne une signification nouvelle à l'ensemble de la théorie. Comme le note Anouk Barberousse, il « jette une lumière nouvelle sur ce qui était simplement considéré comme une conséquence intéressante des lois du mouvement. Le principe de Hamilton, pour ainsi dire, “dit” quelque chose sur la mécanique qui était inconnu avant le travail de Hamilton » (Barberousse, 2008, p. 93).³³

Si l'on suit ce jugement, la formulation lagrangienne actuelle – qui est fondée sur le principe de Hamilton – est bien plutôt un héritage de ce dernier. Cela signifie donc que, contrairement à ce que suggère la distinction que l'on trace aujourd'hui entre mécanique newtonienne d'un côté et mécanique analytique (lagrangienne et hamiltonienne) de l'autre, si l'on prête attention à la signification conceptuelle de la théorie – et non pas seulement à la forme des équations du mouvement – Lagrange se trouve « du côté » de Newton, et non de Hamilton.

Ce constat semble bien confirmer le jugement que Lagrange lui-même émet à pro-

³¹Le concept d'énergie n'apparaît pas en tant que tel chez Hamilton. On le trouve, plus tard, dans les travaux de William Thomson (Kelvin) et de Helmholtz. Cependant, les changements introduits par Hamilton ouvrent la voie à une mécanique de l'énergie. Voir (Darrigol, 2001).

³²Voir (Barberousse, 2008).

³³Anouk Barberousse ajoute, rapportant les propos de Yourgrau et Mandelstam (1968), que, après Hamilton, « la structure conceptuelle de la mécanique a changé d'aspect. [...] La signification, les aspects et les conséquences de la théorie ont changé. » (Barberousse, 2008, p. 93)

pos de sa propre contribution au développement de la mécanique : si les travaux de Hamilton opèrent, un demi-siècle plus tard, un authentique changement conceptuel au sein de la mécanique – au sens défini à la section 1 du présent chapitre –, ceux de Lagrange, en tant que réécriture et confirmation des principes newtoniens, constituent un développement formel, mais non une refonte conceptuelle de la mécanique. En revanche, on peut dire que, si Lagrange avait « raison » de qualifier sa propre contribution de formelle, il avait « tort » de croire que Newton avait « tout dit » et qu'un changement conceptuel était inenvisageable.

À en croire Hamilton (1834) lui-même, si Lagrange ne donne pas au principe de moindre action l'importance fondamentale qu'il acquiert dans le système de Hamilton, c'est parce qu'il n'a pas conçu l'action sous la définition qu'en donne Hamilton³⁴ :

Parce qu'il n'avait pas conçu l'action comme une fonction de cette sorte, les conséquences qui ont été ici déduites de la formule $[\delta V = \sum m(x'\delta x + y'\delta y + z'\delta z) - \sum m(a'\delta a + b'\delta b + c'\delta c) + t\delta H]$ pour la variation de cette intégrale définie semblent avoir échappé à l'attention de Lagrange, et des autres illustres analystes qui ont écrit sur la mécanique théorique, alors même qu'ils étaient en possession d'une formule pour la variation de cette intégrale qui ne différait pas grandement de la nôtre. [...] Il semble, en conséquence, que *c'est avec raison que Lagrange, Laplace et Poisson ont parlé avec légèreté de l'utilité de ce principe dans l'état présent de la dynamique*³⁵. On portera sans doute un jugement différent sur cet autre principe qui a été introduit dans le présent article, sous le

³⁴ « Car bien que Lagrange et d'autres, pour traiter le mouvement d'un système, ont montré que la variation de cette intégrale définie s'annule quand les coordonnées extrémales et la constante H sont données, ils n'ont déduit de ce résultat que la loi bien connue de moindre action; c'est-à-dire, celle selon laquelle si l'on imagine que les points ou les corps composant un système se meuvent depuis un ensemble donné de positions initiales jusqu'à un ensemble donné de positions finales, non pas comme ils le font réellement ni même comme ils pourraient le faire suivant les lois dynamiques générales ou équations différentielles du mouvement, mais de manière à ne violer aucune liaison géométrique supposée, ni cette relation dynamique entre vitesses et configurations qui constitue la loi de la force vive, et si, de plus, on suppose une différence infiniment petite entre ce mouvement géométriquement imaginable, mais dynamiquement impossible, et le mouvement actuel du système entre les positions extrémales données, alors la valeur variée de l'intégrale définie appelée action, ou les forces vives du système accumulées au cours du mouvement ainsi imaginé, différeront infiniment peu de la valeur actuelle de cette intégrale. Mais quand cette loi bien connue de moindre, ou plus précisément d'action stationnaire, est appliquée à la détermination du mouvement actuel du système, elle sert uniquement à former, par les règles du calcul des variations, les équations différentielles de second ordre du mouvement, qui peuvent toujours être trouvées autrement. » (Hamilton, 1834, p. 8)

³⁵ Je souligne. La plupart des successeurs de Lagrange, comme Poisson et Laplace, ne remettent en effet pas en question la centralité du concept de force, et accordent au principe de moindre action une place très secondaire (voir Pulte, 1989, p. 263), comme en témoigne la phrase célèbre de Poisson : « le principe de moindre action est une règle, aujourd'hui inutile, pour construire les équations différentielles du mouvement » (cité dans Pulte, 1989, p. 6).

nom de loi de l'action variée, dans lequel on passe d'un mouvement actuel à un autre mouvement dynamiquement possible, en variant les positions extrémales du système, et (en général) la quantité H , et qui sert à exprimer, par le moyen d'une seule fonction, non seulement les équations différentielles du mouvement, mais leurs intégrales intermédiaires et finales. (Hamilton, 1834, p. 8)

Autrement dit, s'il est « normal » de maintenir le principe de d'Alembert au fondement de la mécanique tant que l'on ne possède pas la définition hamiltonienne de l'action, tout savant post-hamiltonien doit reconnaître le principe de moindre action pour principe fondamental.

Cette lecture des travaux de Hamilton trouve un écho dans les propos de plusieurs historiens, comme Helmut Pulte (1989), qui affirment que les changements introduits par Hamilton constituent une authentique révolution dans l'ontologie sous-jacente de la mécanique et marquent la fin de la mécanique *newtonienne* :

la pensée conservatrice est sans aucun doute complètement étrangère à la philosophie newtonienne de la nature. Le concept d'énergie représente (dans les mots de Kuhn) un nouveau composant fondamental de l'univers et engendre de nouveaux problèmes scientifiques. Si on se demande s'il est encore possible de parler d'une mécanique « newtonienne » dans la seconde moitié du XIX^e, on doit clairement répondre que *non*. (Pulte, 1989, p. 18. Traduit par Maël Pegny).

Ainsi, le changement introduit par Hamilton en mécanique permet le passage d'une théorie de la force à une théorie de l'énergie. Si le newtonianisme – ou d'alembertisme – de Lagrange s'explique par le fait qu'il ne disposait pas du principe de moindre action sous la forme générale que lui a donnée Hamilton, Pulte affirme qu'il n'est plus envisageable, après Hamilton, d'avoir une version lagrangienne (ou newtonienne-d'alembertienne) de la mécanique.

Où en sommes-nous ? Il semble bien que l'on ait, avec les travaux de Lagrange d'une part et de Hamilton d'autre part, les exemples de deux types différents de changements au sein d'un même ensemble logico-empirique : un développement formel pour le premier (sans que cela implique d'en minimiser l'importance), une refonte conceptuelle pour le second. Lagrange maintient en effet, au fondement de sa théorie, un principe newtonien, dont le concept central est celui de force ; en ce sens, il ne fait rien dire de nouveau à la théorie (et le reconnaît). Hamilton, en changeant l'architecture de la théorie, suggère en revanche une interprétation différente : la mécanique newtonienne-d'alembertienne-lagrangienne parle de forces, la mécanique hamiltonienne parle d'énergie. La distinction conceptuelle entre une mécanique vectorielle et une mécanique variationnelle semble donc correspondre à la césure historique entre Lagrange et Hamilton. La mécanique variationnelle d'aujourd'hui, sous ses

formes lagrangienne et hamiltonienne³⁶ correspond en effet à la version de Hamilton et non de Lagrange.

Avant de poursuivre, soulignons une première chose : affirmer ainsi, avec Pulte (1989), que les travaux de Hamilton changent la signification de la mécanique consiste déjà à renoncer à une conception logico-empirique du contenu des théories. Selon cette dernière, en effet, les travaux de Hamilton, tout comme ceux de Lagrange, ne consistent qu'en des modifications formelles, qui n'affectent pas le contenu de la mécanique newtonienne. Instituer une différence entre les contributions de Lagrange et de Hamilton implique déjà que l'on prête attention à l'ordre déductif des principes et non pas seulement à leurs relations logiques « en principe ».

Je voudrais cependant aller plus loin, et montrer que, dès lors que l'on renonce à une définition logico-empirique du contenu des théories et que l'on reconnaît la possibilité d'un changement conceptuel sans modification des conditions de vérité d'une théorie, *il n'est plus même possible d'établir une distinction nette entre développement formel et avancée conceptuelle*. Corrélativement, je souhaite montrer que la césure entre une mécanique de la force et une mécanique de l'énergie ne peut pas être aussi simplement assimilée à un moment historique.

Dans la section suivante, l'examen de plusieurs versions exprimées par des savants, tous post-hamiltoniens, me permettra ainsi de nuancer doublement les conclusions auxquelles semblent conduire l'analyse qui précède. On verra en effet que, contrairement à ce qu'affirme ici Pulte, il est encore possible de préférer une mécanique de la force et d'être lagrangien (ou newtonien) *après* Hamilton. En outre, il est aussi possible de voir dans la contribution de Lagrange – contre l'avis même de ce dernier – une authentique avancée conceptuelle.

³⁶Insistons bien sur le fait que la distinction entre le Lagrange et le Hamilton historiques ne correspond pas à la distinction entre formulations lagrangienne et hamiltonienne d'aujourd'hui : ces dernières sont toutes les deux des formulations variationnelles, post-hamiltoniennes. Leur différence réside dans le type de coordonnées généralisées qu'elles utilisent et dans la définition de l'action, mais pas dans l'ordre général des principes.

vii	
T A B L E.	
<i>PREMIERE PARTIE DE LA MÉCHANIQUE,</i>	
<i>OU LA STATIQUE.</i>	
SECTION I. Sur les différens Principes de la Statique, Page	1
SECT. II. Formule générale pour l'équilibre d'un système quelconque de forces; avec la maniere de faire usage de cette formule.	13
SECT. III. Propriétés générales de l'équilibre déduites de la formule précédente.	25
SECT. IV. Méthode très-simple de trouver les équations nécessaires pour l'équilibre d'un système quelconque de corps regardés comme des points, ou comme des masses finies, & tirés par des puissances données.	44
SECT. V. Solution de différens problèmes de Statique.	58
§. I. De l'équilibre de plusieurs forces appliquées à un même point; & de la composition & décomposition des forces.	<i>Ibid.</i>
§. II. De l'équilibre de plusieurs forces appliquées à un système de corps considérés comme des points, & liés entr'eux par des	

ix	
SECT. IV. Méthode la plus simple pour parvenir aux équations qui déterminent le mouvement d'un système quelconque de corps animés par des forces accélératrices quelconques.	216
SECT. V. Solution de différens problèmes de Dynamique.	233
§. I. Solution générale du problème des oscillations très-petites d'un système quelconque de corps.	241
§. II. Du mouvement d'un corps attiré vers un ou plusieurs centres.	261
§. III. Du mouvement de plusieurs corps qui agissent les uns sur les autres, soit par des forces d'attraction, soit en se tenant par des fils ou par des leviers.	286
SECT. VI. Sur la rotation des Corps.	337
§. I. Formules générales, relatives au mouvement de rotation.	338
§. II. Équations pour le mouvement de rotation d'un corps solide de figure quelconque, animé par des forces quelconques.	372
§. III. Détermination du mouvement d'un corps grave de figure quelconque.	389
SECT. VII. Sur les Principes de l'Hydrodynamique.	418
SECT. VIII. Du Mouvement des Fluides incompressibles.	437
§. I. Équations générales pour le mouvement des Fluides incompressibles.	438

b

viii	
T A B L E.	
fil ou par des verges.	Page 68
§. III. De l'équilibre d'un fil dont tous les points sont tirés par des forces quelconques, & qui est supposé parfaitement flexible ou inflexible, ou élastique, & en même tems extensible ou non.	89
§. IV. De l'équilibre d'un corps solide de grandeur sensible & de figure quelconque, dont tous les points sont tirés par des forces quelconques.	116
SECT. VI. Sur les Principes de l'Hydrostatique.	122
SECT. VII. De l'équilibre des fluides incompressibles.	130
SECT. VIII. De l'équilibre des fluides compressibles & élastiques.	155
SECONDE PARTIE DE LA MÉCHANIQUE,	
OU LA DYNAMIQUE.	
SECTION I. Sur les différens Principes de la Dynamique, Page	158
SECT. II. Formule générale pour le mouvement d'un système de corps, animés par des forces quelconques.	189
SECT. III. Propriétés générales du mouvement déduites de la formule précédente.	198
SECT. IV.	

x	
T A B L E.	
§. II. Du mouvement des Fluides pesans & homogènes dans des vases ou canaux de figure quelconque.	473
SECT. IX. Du mouvement des Fluides compressibles & élastiques.	492

Fin de la Table.

E R R A T A.

Page 16, ligne 11, déterminées; *lif.* indéterminées.
 Page 23, ligne 3, s'étoient joints; *lif.* étoient joints.
 Idem. ligne 21, par la détermination; *lif.* pour la détermination.
 Page 53, ligne 15, équation; *lif.* équations.
 Page 125, ligne 5, je; *lif.* te.
 Page 312, ligne 12, A a'; *lif.* A a".
 Page 324, ligne 20, précession; *lif.* précession.
 Page 424, lig. dern. \sqrt{v} ; *lif.* $\sqrt{v-1}$.
 Page 434, ligne 25, inexactitude; *lif.* incertitude.
 Page 477, ligne 7, intérieure; *lif.* extérieure.
 Page 478, ligne 10, divisées; *lif.* dirigées.
 Idem. ligne 11, $\xi = 90^\circ$; *lif.* $\xi = 90^\circ$.
 Page 504, ligne 7 & suiv.; changez a en a.

FIG. 8 – Table des matières de la *Mécanique analytique* de Lagrange (1788).

4 Versions de la mécanique lagrangienne

Dans cette section, je me propose d'examiner plusieurs versions de la mécanique, telles qu'elles ont été exprimées dans des présentations explicites par différents savants depuis la fin du dix-neuvième siècle, c'est-à-dire à une époque où le développement de la mécanique classique est, pour ainsi dire, achevé. Le premier but de cette section est d'approfondir l'analyse de la notion de version et d'en montrer la fécondité. Je m'efforcerai de souligner le lien, dans chacune des présentations étudiées, entre l'ordre des principes, le statut et la signification qui leur sont attribués, les idéaux scientifiques, les options philosophiques et les préférences théoriques de leur auteur.

En outre, comme on le verra, la plupart de ces présentations s'accompagnent de jugements méta-scientifiques, d'ordre historique ou philosophique, sur la contribution des différents acteurs de l'histoire de la mécanique, ainsi que sur la nature (formelle ou conceptuelle) des différences entre les formulations dont elle peut faire l'objet. Le lien entre ce type de jugements et la version dont ils découlent est au cœur du problème qui m'intéresse ici. L'analyse de ce lien me permettra de nuancer les conclusions de la section précédente, selon lesquelles, d'une part, la contribution de Lagrange à la mécanique est strictement formelle, et, d'autre part, Hamilton provoque une authentique révolution conceptuelle après laquelle il n'est plus possible d'être lagrangien et de maintenir une mécanique de la force.

Il convient de dire un mot, avant de les étudier successivement, de la nature des présentations de la mécanique auxquelles je vais m'intéresser ici. Elles font en effet – à l'exception de la dernière – partie d'un vaste ensemble de réflexions méta-scientifiques et philosophiques de la part de plusieurs savants (physiciens, mathématiciens, physiologistes, chimistes) de la fin du dix-neuvième et du début du vingtième siècles, qu'on appelle parfois « savants philosophes » ou encore « physiciens philosophes ».³⁷ Ce sont, en effet, les derniers savants « complets », à la fois scientifiques et épistémologues, à l'heure où la distinction entre les deux métiers se marque institutionnellement, et où la philosophie des sciences devient une discipline.³⁸

³⁷ Les plus célèbres d'entre eux sont Herman von von Helmholtz (1847), Gustav Kirchhoff (1877), Heinrich Hertz (1894), Ernst Mach (1883), Ludwig Boltzmann (1897) et Henri Poincaré (1902, 1905). Les réflexions de James C. Maxwell (1920) peuvent aussi figurer dans cette liste. Signalons aussi les travaux de Thomson (Lord Kelvin) et Tait (1888), qui constituent la somme la plus complète des connaissances de l'époque en mécanique et une des références principales de la plupart de ces auteurs. Enfin, Duhem (1903b) peut entrer dans cette catégorie des « physiciens philosophes », même si son travail sur la mécanique s'apparente plus à une histoire conceptuelle des sciences.

³⁸ Mach crée en 1895 une chaire de philosophie des sciences à l'université de Vienne, qu'il intitule « histoire et théorie des sciences inductives ». Boltzmann l'y succède en 1902. Sur l'institutionnalisation de la philosophie des sciences comme discipline, voir (Moulines, 2006).

Plusieurs questions relatives à la nature et au but des théories scientifiques sont débattues parmi ces savants, parfois de façon assez violente ; l'un des débats les plus importants est celui qui oppose, entre autres, Duhem, Mach et Kirchhoff à Boltzmann et, dans une certaine mesure, Hertz, les premiers affirmant qu'une théorie doit être une représentation ou description simple et économique des phénomènes qui en permet la prédiction, quand les seconds défendent l'idée qu'une théorie, en sus de prédire, doit expliquer les phénomènes.³⁹ Le débat est d'autant plus complexe que s'affrontent plusieurs conceptions de la représentation que doit fournir une théorie, conceptions sous-tendues par des « visions du monde » différentes ; ainsi, à une image du monde mécaniste et atomiste, défendue par exemple par Boltzmann, Maxwell et Lord Kelvin (William Thomson), Mach oppose une conception phénoméniste, qui invite à s'en tenir aux seuls phénomènes sans prétendre décrire la réalité qui en serait la cause, et Duhem défend l'énergétisme, doctrine selon laquelle la quantité physique fondamentale est l'énergie.

L'analyse de ces débats dépasserait les limites du présent travail, et n'y a pas sa place. Je me contenterai, dans ce qui suit, d'examiner certains aspects des versions de la mécanique exprimées dans les présentations de trois d'entre ces savants : Ernst Mach (1883), Heinrich Hertz (1894) et Pierre Duhem (1903a,b). C'est cependant parce qu'elles prennent sens dans le cadre de ces débats complexes que ces présentations sont particulièrement intéressantes pour mon propos. La mécanique étant encore, à cette époque, considérée comme *la* théorie physique par excellence et comme un modèle auquel les autres branches de la physique doivent s'efforcer de se conformer, il est naturel de s'y référer quand on réfléchit sur la nature des théories scientifiques. Chacun y cherche une confirmation de ses convictions philosophiques, et dans le même mouvement expose la mécanique telle qu'il la comprend, à la lumière desdites convictions. De là naissent de nombreux désaccords sur la signification des différents principes, sur la présentation qu'il convient d'en faire, et sur la place de la mécanique dans l'architecture de la connaissance scientifique. Comme le remarque Hertz, « tous les physiciens s'accordent pour dire que la tâche de la physique consiste à rapporter les phénomènes de la nature aux lois simples de la mécanique. Mais quand il s'agit de savoir quelles sont ces lois simples, il ne règne plus la même unanimité. » (Hertz, 1894, p. 63)

Ainsi, les différentes versions que je me propose d'étudier se présentent comme des

³⁹Cette opposition ne recoupe pas celle qui oppose réalistes et instrumentalistes. On peut, comme Boltzmann, attendre d'une théorie qu'elle fournisse une explication des phénomènes à l'aide d'une analogie mécanique éclairante sans supposer que cette théorie est vraie. Pour une analyse des notions d'explication et de compréhension dans les travaux de Boltzmann, voir (de Regt, 1999). Voir aussi (de Courtenay, 1999; Uffink, 2004).

manières *concurrentes* de comprendre la mécanique. Contrairement à ce qui est le cas dans l'analyse du chapitre 2, où la coexistence de plusieurs formulations standard ne pose pas de problème particulier aux praticiens et aux théoriciens de la mécanique⁴⁰ et n'implique aucunement la nécessité d'effectuer un *choix* entre elles (voir page 81), les différentes versions de la mécanique exprimées par les présentations de ces auteurs entrent – parfois explicitement – en conflit les unes avec les autres. Dans bien des cas, c'est même la recherche de la « meilleure » présentation des principes de la mécanique, celle qui en exprimerait la version la plus limpide, cohérente et dépourvue d'obscurité, qui motive l'entreprise de l'auteur.

Signalons en effet que, si les options philosophiques des différents savants prenant part au débat sur les principes de la mécanique sont opposées, chacun se réclame d'un idéal de simplicité et, pour reprendre les termes de Mach, d'un certain « esprit antimétaphysique » : il s'agit de dépouiller le plus possible les énoncés de la mécanique de concepts dont la signification est douteuse. Chacun, ainsi, cherche à présenter la mécanique de la manière la plus simple et *intelligible* possible. Le problème est que chacun entend ces termes de « simplicité », de « métaphysique » et d'« intelligibilité » d'une manière qui traduit des convictions, explicites et parfois implicites, profondément différentes. Ainsi, comme le dit Hertz,

la question de savoir ce qui est ici encore simple et admissible, et ce qui ne l'est déjà plus, n'est pas tranchée ; c'est précisément sur ce point que l'on cesse d'invoquer l'unanimité générale. C'est pourquoi nous voyons effectivement naître de réelles divergences d'opinion quant à savoir si telle ou telle hypothèse correspond encore ou non à la mécanique ordinaire. [...] Ainsi, par exemple, est-il prématuré de chercher à rapporter les équations du mouvement de l'éther aux lois de la mécanique tant que l'on ne s'est pas entendu de façon univoque sur ce que l'on prétend avoir désigné par ce nom. (Hertz, 1894, p. 63)

Je me propose donc, dans ce qui suit, d'examiner successivement les présentations de la mécanique de Mach (1883), de Hertz (1894), et de Duhem (1903b). Je les considérerai à la fois comme l'expression d'autant de versions de la mécanique apportant du matériau historique à mon analyse, et comme celle des réflexions de philosophes confrontés au problème même qui m'a conduite à introduire la notion de version : on trouvera en effet, sous la plume de chacun de ces auteurs, une analyse de la relation entre le contenu de la mécanique, la forme sous laquelle elle est présentée, et la compréhension que les agents individuels, en fonction de leur histoire et de

⁴⁰ Les philosophes des sciences s'y sont d'ailleurs eux-mêmes très peu intéressés ; mis à part (Butterfield, 2004) et (North, 2009a,b), il est rare de trouver des analyses approfondies de ce cas, qui se trouve plutôt convoqué comme la preuve qu'une théorie n'est pas réductible à l'une de ses formulations, preuve reposant elle-même sur le postulat selon lequel ces différentes formulations sont équivalentes.

leur psychologie propre, en ont. Je terminerai cette section par un rapide examen de la version de la mécanique exprimée par Cornelius Lanczos (1970), mathématicien et physicien hongrois et auteur d'un des ouvrages aujourd'hui considéré comme une référence en mécanique analytique, qui m'offrira donc l'exemple d'une version de la mécanique assez largement partagée par les physiciens d'aujourd'hui.

4.1 Ernst Mach (1883), *La Mécanique. Exposé historique et critique de son développement*

Ernst Mach écrit son « exposé historique et critique » du développement de la mécanique un siècle après la parution de la *Mécanique analytique* de Lagrange, et un demi-siècle après les travaux de Hamilton. Comme on va le voir, sa version de la mécanique, semblable en de nombreux points à celle de Lagrange, offre un remarquable exemple de la possibilité, pour un physicien post-hamiltonien, de choisir une mécanique de la force.

4.1.1 Ancrer la signification des principes dans une expérience sensible : la défense du concept de force

Exposer rigoureusement la méthode mise en œuvre par Mach nécessiterait une analyse approfondie de sa philosophie des sciences et de ses travaux en physiologie de la perception⁴¹ et m'éloignerait déraisonnablement de mon propos. En quelques mots, disons que Mach défend une position phénoméniste radicale, selon laquelle il n'existe aucun monde physique au delà des apparences, c'est-à-dire de nos expériences sensibles. La science doit viser à fournir des formules mathématiques permettant de décrire les objets de nos expériences sensibles de la manière la plus simple et la plus économique, sans affirmer que ces formules sont l'expression de lois de la nature. Son entreprise d'examen historique et critique des principes de la mécanique vise ainsi à retrouver, pour chacun des principes utilisés pour représenter et prédire les phénomènes du mouvement, sa *signification véritable*, en le réduisant au constat de ce qu'il appelle un « fait » empirique familier, c'est-à-dire d'une expérience sensible, réelle ou possible. Ainsi par exemple, à propos du principe des déplacements virtuels :

Il est important de bien prendre conscience que [...] nous n'avons affaire qu'au constat et à l'établissement d'un *fait*. [...] Notons [...] que le principe des déplacements virtuels, comme tout principe général, apporte, par l'intuition qu'il nous offre, *désillusion* et clarté. Il apporte de la désillusion dans la mesure où

⁴¹Mach les présentera quelques années plus tard dans l'*Analyse des sensations* (Mach, 1886). Pour une analyse de la philosophie de Mach, voir (Cohen et Seeger, 1970; Blackmore, 1972; Banks, 2003).

nous y reconnaissons des faits connus depuis longtemps et même perçus instinctivement ; en les reconnaissant, nous en avons une perception plus distincte et définie. Il apporte de la clarté, dans la mesure où il nous permet de toujours voir ces mêmes faits simples, jusque dans les relations les plus compliquées. (Mach, 1883, pp. 87-88)⁴²

En reconstituant les expériences qui ont eu lieu ou auraient pu avoir lieu, Mach reconstruit le processus qui a pu conduire à la formulation des différents principes. Par cet exposé, il cherche à rendre les différents principes de la mécanique « intuitifs » et, partant, intelligibles.⁴³

Mach insiste à plusieurs reprises sur le fait que le « goût scientifique » (p. 307) et la « nature intellectuelle »⁴⁴ des savants jouent un rôle important dans le développement historique de la mécanique et dans la manière et l'ordre choisis par chacun pour exposer les principes. Ainsi note-t-il, par exemple, que tout au long du développement de la mécanique, l'attention des chercheurs a été attirée par des « aspects très différents des phénomènes mécaniques » : on considère tantôt que « la quantité de mouvement d'un système [est] déterminée par la force », tantôt que « sa *force vive* [est] déterminée par le travail ». Et Mach ajoute que « le caractère propre des chercheurs joue un grand rôle dans le choix de l'aspect » qui retiendra son attention (Mach, 1883, p. 315).

Mach reconnaît donc clairement la variabilité de ce que j'ai proposé d'appeler les « versions » de la mécanique et souligne leur importance dans le développement de la science. De plus, loin de chercher à débarrasser la mécanique de ce qu'elle doit à son ancrage dans la sensibilité humaine, tout idiosyncratique soit-elle, Mach considère qu'un principe n'est intelligible qu'en tant qu'il trouve un écho dans notre sensibilité. Tant qu'elle ne donne pas lieu à des contradictions, il est donc, selon lui, tout à fait légitime de faire appel à la signification intuitive des concepts scientifiques. Le pari de Mach est précisément de fonder toute la mécanique sur une intuition raisonnée.

⁴²Toutes les citations de Mach sont traduites par moi-même à partir de la traduction anglaise indiquée en bibliographie.

⁴³Un exemple, parmi d'autres, de cette reconstruction des principes sur une base intuitive est offert par l'exposé de la démonstration par Lagrange du principe des déplacements virtuels s'appuyant sur la fiction d'un train de poulies et le commentaire qu'il lui adjoint : « La déduction de Lagrange, si on la dépouille de la fiction quelque peu étrange des poulies, a quelque chose de vraiment convaincant, en raison du fait que l'action d'un poids unique est bien plus proche de notre expérience et plus aisée à suivre que l'action de plusieurs poids. » (Mach, 1883, p. 78). Voir aussi le diagramme (reproduit au chapitre 7 du présent chapitre, figure 31, page 385) destiné à faire comprendre le principe de Maupertuis, qui met en évidence la signification mathématique d'un principe de minimum (Mach, 1883, p. 81).

⁴⁴Voir ses considérations sur les natures intellectuelles respectives de Galilée et de Kepler (Mach, 1883, pp. 307-310).

En conséquence, cette conviction philosophique a une influence sur la version de Mach lui-même. En effet, il propose une défense du concept de force au nom de ce qu'il a d'intuitif et même de pré-scientifique. C'est donc en grande partie en raison de sa volonté d'ancrer les principes de la mécanique dans des expériences sensibles familières qu'il place le concept de force au centre de sa version, considérant que la mécanique en tire sa signification. Ainsi écrit-il, à propos de l'analogie entre le concept de force et l'action humaine :

Les tentatives de rejeter cette conception comme subjective, animiste et non scientifique sont invariablement vouées à l'échec. En effet, il n'est d'aucune utilité de faire violence aux pensées qui nous sont naturelles et de nous condamner ainsi à une volontaire misère intellectuelle. (Mach, 1883, p. 95)

Le concept mécanique de force, généralement représenté par un vecteur, trouve une signification intuitive dans notre propre expérience d'êtres agissants soumis aux lois de la mécanique :

Ces deux principes⁴⁵ contiennent [...] *un énoncé généralisé de la loi de l'inertie*, dont, sous cette forme, non seulement nous *voyons*, mais encore nous *sentons* la justesse.

Ce sentiment n'est pas anti-scientifique ; il est encore moins nuisible. Sans se substituer à la compréhension conceptuelle, il l'accompagne et il est la condition fondamentale et la seule preuve d'une maîtrise *complète* des faits mécaniques. Nous sommes nous-mêmes des systèmes mécaniques, et cela modifie profondément notre vie mentale. Personne ne pourra nous convaincre que la considération des processus mécanico-physiologiques, ainsi que des sentiments et des instincts qu'ils impliquent, doive être exclue de la mécanique scientifique. Si l'on ne connaît des principes comme ceux du centre de gravité et des aires que sous leur forme mathématique abstraite, sans avoir eu affaire aux simples faits palpables, qui sont à la fois leur application et leur source, on ne les comprend qu'à moitié, et l'on aura du mal à reconnaître les phénomènes réels comme des exemples de la théorie. On est alors dans la situation d'une personne qui, sans avoir parcouru la région environnante, se trouverait soudain en haut d'une tour, et serait incapable de reconnaître les objets qu'elle aurait sous les yeux. (Mach, 1883, pp. 394-395)

Sur ce point, une comparaison avec le jugement que Poincaré porte sur ce qu'il appelle la « mécanique anthropomorphique » permet de mettre en évidence la manière dont les conceptions philosophiques et les idéaux scientifiques de chacun à la fois fondent leur version de la mécanique et rendent les conflits entre versions difficiles à trancher. Poincaré prend en effet l'exact contre-pied de Mach, en soulignant le

⁴⁵Il s'agit du théorème du centre de gravité et du théorème des aires.

caractère illégitime et vain d'un appel à l'aspect intuitif du concept de force pour le placer au fondement de la théorie :

l'idée de force est une notion primitive, irréductible, indéfinissable ; nous savons tous ce que c'est, nous en avons tous l'intuition directe. Cette intuition directe provient de la notion d'effort, qui nous est familière depuis l'enfance.

Mais d'abord, quand même cette intuition directe nous ferait connaître la véritable nature de la force en soi, elle serait insuffisante pour fonder la Mécanique ; elle serait d'ailleurs tout à fait inutile. Ce qui importe, ce n'est pas de savoir ce que c'est que la force, c'est de savoir la mesurer.

Tout ce qui ne nous apprend pas à la mesurer est aussi inutile au mécanicien, que l'est, par exemple, la notion subjective de chaud et de froid au physicien qui étudie la chaleur. [...]

Mais il y a plus : cette notion d'effort ne nous fait pas connaître la véritable nature de la force ; elle se réduit en définitive à un souvenir de sensations musculaires, et on ne soutiendra pas que le soleil éprouve une sensation musculaire quand il attire la terre.

Tout ce qu'on peut y chercher, c'est un symbole, moins précis et moins commode que les flèches dont se servent les géomètres, mais tout aussi éloigné de la réalité. L'anthropomorphisme a joué un rôle historique considérable dans la genèse de la mécanique ; peut-être fournira-t-il encore quelquefois un symbole qui paraîtra commode à quelques esprits ; mais il ne peut rien fonder qui ait un caractère vraiment scientifique, ou un caractère vraiment philosophique. (Poincaré, 1902, pp. 124-125)

Il est important de préciser, toutefois, que la défense du concept de force dans ce qu'il a d'anthropomorphique est clairement distincte, chez Mach, de l'adoption d'une ontologie de la force. La science, selon Mach, ne parle que des apparences sensibles ; en conséquence, il n'y a, dans le monde, ni force, ni atomes, ni énergie. Selon la sensibilité de chacun, on peut, comme Newton, préférer parler de force, ou, comme Huyghens, choisir le travail comme concept premier ; on organisera alors différemment les principes de la mécanique : « On est donc libre de considérer la *durée* de la chute ou la *distance* parcourue comme le paramètre déterminant de la vitesse. Si l'on centre l'attention sur la première, le concept premier est celui de force, et le travail est un concept dérivé ; si, en revanche, on étudie l'influence de la seconde, alors le concept de travail est premier. » (Mach, 1883, p. 311)

Ces deux versions, fondées sur des expériences sensibles différentes, sont cependant également légitimes et rien ne permet, selon Mach, de trancher en faveur de l'une ou l'autre.

Après Huyghens et Newton, les deux manières de penser se mêlèrent et, leur équivalence et leur indépendance n'étant pas toujours reconnues, cela conduisit

à plusieurs erreurs et confusions, en particulier dans la querelle entre les cartésiens et les leibniziens [...] à propos de la mesure de la force. Plus récemment, cependant, les chercheurs préfèrent tantôt l'une, tantôt l'autre manière de penser. [...] Cependant, si l'on y regarde de près, on s'aperçoit que les principes de Huyghens et les propositions de Newton sont également simples et évidents. (Mach, 1883, pp. 311-314)

Ainsi, la diversité des versions de la mécanique est celle des expériences sensibles de différents individus, mais ne doit pas être comprise comme la mise en concurrence de plusieurs interprétations physiques différentes. Elle ne remet pas en cause, pour Mach, l'identité profonde de la théorie et, comme on va le voir, l'équivalence des différents principes au moyen desquels elle peut être exprimée.

4.1.2 Distinguer entre le contenu physique et la présentation formelle de la mécanique : intelligibilité des principes et économie de la pensée

La recherche d'un ancrage de la signification physique des principes dans une expérience sensible, jointe à l'exigence d'« économie de la pensée »⁴⁶, qu'il considère comme la vertu suprême d'une théorie scientifique, conduit Mach à distinguer clairement entre la signification ou le *contenu physique* de la mécanique, que l'enquête historique et critique permet de retrouver et de rendre intelligible, et sa *présentation formelle*, remarquable accomplissement de la science comme économie de la pensée, dont la contrepartie est cependant de nous rendre les principes moins intelligibles, parce que plus abstraits.

Son jugement sur la signification des différents principes de la mécanique est sans équivoque : ils sont tous équivalents et se réduisent à la constatation d'une même catégorie de faits empiriques⁴⁷, intégralement exprimés dans les principes newtoniens, mais que les mathématiciens postérieurs à Newton ont contribué à résumer dans des formules plus générales, commodées et économiques :

La contribution de Newton au sujet qui nous intéresse ici est double. Premièrement, il a considérablement étendu le domaine de la mécanique physique en découvrant la *gravitation universelle*. Deuxièmement, il a *achevé l'énoncé formel des principes de la mécanique tels qu'ils sont aujourd'hui généralement acceptés*. Depuis Newton, aucun principe essentiellement nouveau n'a été énoncé. Tout ce qui a été accompli en mécanique depuis lors a consisté en un développement

⁴⁶ « La science elle-même peut [...] être considérée comme un problème de minimum, qui consiste à présenter les faits de la façon la plus complète possible avec *la moindre dépense intellectuelle*. » (Mach, 1883, p. 586)

⁴⁷ « Ainsi, malgré toutes leurs différences apparentes, les principes de la mécanique ont une identité profonde. Ces principes ne sont pas les expressions de différents faits, mais, en quelque sorte, les reflets de différents *aspects* du même fait. » (Mach, 1883, pp. 478-479)

déductif, formel, et mathématique de la mécanique à partir des lois de Newton.
(Mach, 1883, p. 226)

À eux seuls, les principes de Newton suffisent, sans qu'il soit nécessaire d'introduire de nouvelles lois, à explorer entièrement tous les phénomènes mécaniques que l'on peut rencontrer en pratique, que ce soit en statique ou en dynamique. Les difficultés qui se présentent dans ces problèmes sont invariablement d'un caractère mathématique (formel), et ne concernent en aucun cas les principes.
(Mach, 1883, p. 344)

Non seulement, selon Mach, le développement formel dont la mécanique a fait l'objet depuis Newton ne modifie aucunement le contenu physique de la théorie, mais encore il se distingue nettement de ce que Mach appelle la « période déductive » de l'histoire de la théorie. Il est le signe d'une science arrivée à maturité, pour laquelle on est en possession de tous les principes et théorèmes qui en constituent le contenu :

Une fois que les faits les plus importants d'une science de la nature ont été établis par l'observation, son développement entre dans une nouvelle période – la période *déductive* [...]. Au cours de cette période, on peut reproduire mentalement ces faits sans avoir constamment recours à l'observation. On se représente, par la pensée, les cas les plus généraux et compliqués, comme composés d'éléments plus simples et familiers de l'observation. Cependant, même après que l'on a déduit, de l'expression des faits les plus élémentaires (les principes) l'expression des faits plus communs et compliqués (les théorèmes) et que l'on a découvert, dans tous les phénomènes, les mêmes éléments, le développement de la science n'est pas achevé. Le développement déductif de la science est suivi de son développement *formel*. Il s'agit alors de présenter sous une forme claire et concise, c'est-à-dire sous la forme d'un *système*, les faits qu'il s'agit de reproduire, de telle sorte que chacun d'entre eux puisse être atteint et représenté mentalement avec le *moindre effort intellectuel*. (Mach, 1883, p. 516)

Il est tentant, ici, de reprendre la distinction suggestive de Feynman entre physicien et mathématicien (voir page 114), et d'affirmer que le travail déductif, qui consiste à mener une enquête systématique dans les conséquences physiques de la théorie et à chercher la signification empirique de chacun des théorèmes que l'on peut déduire des principes généraux, est typiquement celui du physicien, quand le développement formel est l'œuvre du mathématicien.

Or, la mécanique analytique, qu'il définit comme une « mécanique *algébrique* », par opposition à la mécanique de Newton, qui est « purement *géométrique* » (Mach, 1883, pp. 560-561), est, à ses yeux, un des exemples les plus remarquables que l'histoire des sciences offre du *développement formel d'une science* :

C'est Lagrange qui a porté la mécanique analytique à son plus haut degré de perfection. Le but de Lagrange, dans sa *Mécanique analytique* (1788), était de se débarrasser, *une fois pour toutes*, de tous les raisonnements nécessaires pour

résoudre des problèmes mécaniques, en les réduisant autant que possible à une seule et unique formule. Il y est parvenu. Tous les cas susceptibles de se présenter peuvent désormais être traités en suivant un schéma simple, hautement symétrique et clair, et tous les raisonnements qui restent sont effectués par des calculs purement mécaniques. La mécanique de Lagrange est une contribution prodigieuse à l'économie de la pensée. (Mach, 1883, pp. 561-562)

Mais ce travail est clairement distinct de celui qui consiste à enquêter sur la signification physique des principes :

Il ne faut pas attendre de cette branche de la mécanique des éclaircissements de principe. Au contraire, la connaissance des principes doit être essentiellement achevée avant que l'on puisse envisager d'élaborer une mécanique analytique ; le seul but de cette dernière est une *maîtrise* pratique parfaite des problèmes. Si l'on n'est pas conscient de cela, on ne peut pas comprendre l'importante contribution de Lagrange, qui est, ici aussi, essentiellement *économique*. (Mach, 1883, p. 575)

On a ici la preuve que l'affirmation de Pulte (page 119), selon laquelle la mécanique de la deuxième moitié du dix-neuvième siècle ne peut pas être une mécanique de la force, est à nuancer : alors que Mach connaît le principe de moindre action sous sa forme hamiltonienne, il n'est pas question pour lui d'y voir un principe nouveau. Tous les principes de minimum sont équivalents, affirme-t-il à plusieurs reprises : « on peut envisager *de nombreuses autres* expressions intégrales qui donnent, par variation, les équations ordinaires du mouvement, sans que pour autant elles aient une signification *physique* particulière. » (Mach, 1883, p. 463)⁴⁸ Ils sont, de plus, équivalents au principe de d'Alembert : « On déduit aisément le principe de Hamilton de celui de d'Alembert et, inversement, celui de d'Alembert à partir de celui de Hamilton ; ces deux principes sont en fait identiques, et leurs différences ne sont que formelles. » (Mach, 1883, p. 475)

L'analyse de ces propos de Mach invite donc à établir une distinction entre deux types de « principes premiers » : on peut très bien reconnaître qu'une certaine formule, comme le principe de Hamilton, est la plus économique de toutes celles dont on dispose, et la placer légitimement au fondement de l'édifice formel de la mécanique, sans pour autant considérer que c'est de cette formule que la théorie tire sa signification physique.

Plus précisément, en vertu du fait que tous les principes différentiels et intégraux de la mécanique ne sont, selon Mach, que des formes différentes du principe fondamental donné par Newton, on peut indifféremment choisir l'un ou l'autre comme

⁴⁸Voir aussi (Mach, 1883, p. 475) : « On a remarqué, dans ce qui précède, que l'on peut proposer *plusieurs* expressions telle que l'annulation de leur variation conduise aux équations ordinaires du mouvement. Le principe de Hamilton contient une expression de ce genre. »

principe premier. Quelle que soit l'allure que prendra l'édifice de la mécanique, sa signification en restera inchangée. Or, en ce qui concerne la présentation formelle, c'est l'exigence d'économie et de commodité qui préside :

La mécanique n'a [...] pas seulement sa fin en elle-même; elle doit aussi *résoudre des problèmes* pour des besoins pratiques et pour développer les autres sciences. Ces problèmes sont pour la plupart plus faciles à résoudre par d'autres méthodes que celles de Newton – des méthodes dont il a été prouvé qu'elles sont équivalentes à ces dernières. En conséquence, il serait vain et coûteux de dédaigner leurs avantages et d'insister toujours pour que l'on revienne aux idées newtoniennes élémentaires. Il est suffisant de s'être convaincu une fois que ce retour est toujours possible. Ajoutons cependant que les conceptions newtoniennes sont certainement les plus *satisfaisantes* et les plus claires [...]. (Mach, 1883, pp. 356-357)

Pour résoudre des problèmes de mécanique, le principe de d'Alembert s'avère d'une grande utilité. C'est un résultat remarquable du développement formel de la mécanique, orienté par un idéal d'économie. Cependant, en raison même de cette exigence d'économie intellectuelle, les savants sont parfois conduits à faire usage de formules aveugles, dont la signification provient d'un principe dont elles sont une reformulation commode. C'est précisément le travail de l'histoire critique faite par Mach que de retrouver la signification physique des différents principes, pour nous assurer que l'on peut continuer à faire usage du principe de d'Alembert et des principes de minimum qui lui sont équivalents⁴⁹ :

Le principe de d'Alembert ne nous apprend rien [...] que nous ne puissions apprendre par d'autres méthodes. Le principe remplit, au cours de la résolution des problèmes, la fonction d'un gabarit qui nous épargne, dans une certaine mesure, d'avoir à examiner chaque nouveau cas, en nous offrant une méthode permettant d'utiliser des expériences connues et familières. Le principe ne nous offre pas tant une *vue* plus profonde des processus étudiés qu'une *maîtrise pratique*. La valeur de ce principe est d'ordre économique.

Quand on a résolu un problème au moyen du principe de d'Alembert, on peut se contenter de savoir que l'application de ce principe implique toutes les expériences concernant l'équilibre que l'on a faites auparavant. Cependant, si l'on veut appréhender *clairement et complètement* le phénomène, c'est-à-dire y retrouver les éléments mécaniques les plus simples qui nous sont familiers, on est

⁴⁹ « La résolution de quelques problèmes au moyen du principe de d'Alembert suffit à montrer combien il est commode. Elle nous donne aussi la conviction qu'il est possible, dans tous les cas où cela semblera nécessaire, de résoudre directement et en toute clarté le même problème par la considération des processus mécaniques élémentaires, et de parvenir au même résultat. La conviction de la *possibilité* de cette opération en rend la réalisation inutile dans les cas où le seul but poursuivi est pratique. » (Mach, 1883, pp. 432-433)

obligé de prolonger notre étude et de remplacer les expériences concernant l'équilibre par les conceptions de Newton ou de Huyghens.⁵⁰ (Mach, 1883, pp. 430-431)

Le principe de d'Alembert n'est donc, selon Mach, qu'un résumé d'expériences, qui ne donne en aucun cas une compréhension des phénomènes, mais qui permet à celui qui a compris une fois les raisons de sa validité, de faire l'économie de certains raisonnements. Ce principe ne dit rien par lui-même⁵¹, et ne nous apprend rien de plus que ce que nous disent les principes newtoniens.

Pour résumer, tous les principes de la mécanique sont, selon Mach – comme, d'ailleurs, selon Lagrange – équivalents, les développements formels obéissant à un idéal d'économie qui ne change rien à l'ancrage empirique de la théorie. Mais cet ancrage lui-même, qui donne sa signification physique à la théorie, est susceptible d'être appréhendé de différentes manières, en fonction de la tournure d'esprit de chacun (voir ci-dessus, page 129).

En distinguant nettement entre le principe premier du point de vue de la généralité et de l'économie, et le principe premier du point de vue de la signification physique, Mach prouve qu'il est possible de rester lagrangien (ou newtonien) après Hamilton : même une fois que le principe de moindre action a une forme suffisamment générale pour pouvoir figurer au fondement de l'édifice formel de la mécanique, le contenu de cette dernière dépend des expériences sensibles qui lui donnent un sens.

Ici encore, la question de savoir ce qui, chez Mach, est déterminant et premier, d'un engagement inavoué en faveur des concepts newtoniens (en particulier celui de force), de sa méfiance vis-à-vis de tout principe susceptible d'appeler une interprétation métaphysique, ou encore de la conception philosophique des théories scientifiques qu'il défend et qui le conduit à distinguer nettement entre la signification physique des principes et leur exposé formel, n'est pas de celles que je cherche à résoudre. J'espère en revanche avoir mis en évidence, par cet examen, le caractère indissociable de ces différents déterminants de ce que j'appelle la version d'un individu. D'autre part, l'analyse de la manière dont Mach traite la question de la signification physique d'une théorie et de la distinction entre le conceptuel et le formel me fournit un premier élément, auquel vont s'ajouter les analyses des versions de Hertz et de Duhem, permettant de montrer que cette distinction même dépend de la version de celui qui l'établit.

⁵⁰Voir ci-dessus, page 129.

⁵¹Selon Mach, on ne peut donc pas, à proprement parler, *raisonner* au moyen du principe de d'Alembert, au sens où le raisonnement impliquerait la manipulation d'un contenu conceptuel. Le principe de d'Alembert ne peut faire l'objet que de ce que Leibniz aurait appelé une *cogitatio caeca* (voir Leibniz, 1765, II, XXI, p. 146).

4.2 Heinrich Hertz (1894), *Les Principes de la mécanique*

Heinrich Hertz partage beaucoup des exigences philosophiques de Mach, en particulier son idéal d'économie ou de commodité. Il affirme d'ailleurs, dans la préface de ses *Principes*, devoir « énormément au beau livre de Mach sur le développement de la mécanique. » (Hertz, 1894, p. 6) Mach, à son tour, dans une édition ultérieure de sa *Mécanique*, consacre une section à l'analyse du livre posthume de Hertz, et reconnaît que « l'exigence de commodité de Hertz concorde avec [sa propre] exigence de simplicité. » (Mach, 1883, p. 250)⁵² Tout en se réclamant du même « esprit anti-métaphysique », Hertz expose cependant une version de la mécanique qui se démarque nettement, et sur de nombreux points, de celle de Mach : comme on va le voir, les différents principes de la mécanique ne sont pas, pour Hertz, équivalents ; les deux grands types de présentations qui en ont été faites (mécanique de la force et mécanique de l'énergie) n'expriment pas le même contenu. D'autre part, il critique vigoureusement le concept de force.⁵³ Avant d'examiner sa version proprement dite, je propose de revenir rapidement sur la conception philosophique des théories qui guide l'entreprise de Hertz.

4.2.1 Images, versions et présentations

L'exposé de Hertz (1894) est peut-être le meilleur exemple d'un examen rationnel des principes de la mécanique fait au nom d'une certaine conception des théories scientifiques. L'introduction de son ouvrage est ainsi consacrée à l'exposé de sa célèbre conception des théories comme « images ». L'idée fondamentale des *Bildtheorien*, telles que Hertz, mais aussi Boltzmann (1905)⁵⁴ les défendent, est que l'activité théorique consiste en la construction mentale d'une image ou modèle des phénomènes :

[...] le procédé que nous utilisons invariablement pour dériver le futur à partir du passé et parvenir ainsi à la prévision souhaitée est le suivant : nous fabriquons en nous des simulacres, ou des symboles internes des objets externes, et nous les fabriquons de telle sorte que les conséquences, nécessaires selon la pensée, de ces images soient toujours les images des conséquences, nécessaires selon la

⁵²Pour une comparaison entre les vues de Hertz et celles de Mach, voir par exemple (Preston, 2008).

⁵³Pour une analyse récente des *Principes de la mécanique* de Hertz, voir (Lützen, 2005).

⁵⁴La question de savoir qui, de Hertz et de Boltzmann, est le premier à avoir thématiqué cette idée, n'est pas tranchée (voir de Regt, 1999, p. 117). Pour une comparaison des positions de Hertz et de Boltzmann, ainsi que de leur influence sur Wittgenstein, voir (Wilson, 1989; D'Agostino, 1990; Bouveresse, 1991). Sur la *Bildtheorie* et la philosophie de Boltzmann, voir aussi (Blackmore, 1995; de Regt, 1999; Visser, 1999), ainsi que l'ensemble du numéro spécial de la revue *Synthese* d'avril 1999, vol. 119, n°1 et 2.

nature, des objets reproduits. [...] Une fois que nous avons réussi, à partir de l'expérience déjà accumulée, à dériver des images qui soient du type exigé, nous pouvons en peu de temps développer à partir d'elles, comme à partir de modèles réduits, les conséquences qui, dans le monde extérieur, n'auront lieu qu'au bout d'un certain temps ou qui résulteront de notre propre intervention ; [...]. (Hertz, 1894, p. 67)⁵⁵

Sans qu'il soit question de l'examiner en détails ici, notons que cette conception des théories défendue par Hertz peut être interprétée comme une opposition frontale à la philosophie d'un Mach ou d'un Duhem qui ne voient dans les théories que des ensembles de symboles abstraits commodes. Hertz ne se prononce pas explicitement sur la question de l'explication, mais Boltzmann (1905) ancre dans sa *Bildtheorie* l'affirmation selon laquelle une théorie doit fournir une explication des phénomènes et n'est explicative qu'en tant qu'elle en offre une image intelligible, c'est-à-dire une image conforme aux « lois de notre pensée », de telle sorte que nous soyons capables de la manipuler.⁵⁶

On peut légitimement voir, dans la position de Hertz, et plus encore dans celle de Boltzmann, l'expression de la conviction qui fonde l'ensemble du présent travail, et que j'ai exposée dès l'introduction générale : une théorie ne remplit véritablement sa double fonction d'outil de représentation et d'inférence qu'en tant qu'elle est comprise par des agents. Autrement dit, c'est au moyen des représentations mentales qui constituent ma version de la mécanique – des « images », dans les termes de Hertz et de Boltzmann – que je parviens à me représenter les phénomènes et, en manipulant ces représentations, à les prédire.

Affirmer ainsi le caractère indissociable de la fonction de représentation et de la fonction inférentielle des théories – et, corrélativement, l'exigence selon laquelle une théorie doit être explicative – ne revient cependant pas à adopter *ipso facto* une position réaliste. Bien au contraire, admettre que les théories sont essentiellement d'« étoffe mentale » invite à accepter que l'idéal d'objectivité poursuivi par l'enquête scientifique ne pourra jamais être tout à fait atteint. Certes, pour permettre la prédiction des phénomènes, nos images doivent mettre en évidence, comme le dit également Poincaré (1905, p. 175), un « rapport vrai », mais elles comportent aussi inévitablement des éléments qui n'ont pas de contrepartie dans le monde physique :

⁵⁵De même, Boltzmann écrit que « la tâche de la théorie est de construire une image du monde extérieur qui existe de façon purement interne et qui doit être le guide de nos pensées et de nos expériences » (extrait d'un discours de 1890 intitulé « Sur la signification des théories », reproduit dans Boltzmann, 1974, p. 33).

⁵⁶Voir (de Regt, 1999). Il existe un désaccord important entre Hertz et Boltzmann sur la nature même des « lois de la pensée ». Je le présenterai au chapitre 4, section 2.1.

Les images dont nous parlons sont nos représentations des choses; elles possèdent précisément avec les choses *cette* concordance essentielle qui consiste dans la satisfaction de l'exigence en question; mais pour atteindre leur but, il n'est nullement nécessaire que ces images présentent avec les choses quelque autre concordance que ce soit. De fait, nous ne savons pas non plus si nos représentations des choses concordent avec elles en quoi que ce soit d'autre que, précisément, *cette* relation fondamentale – et nous n'avons aucun moyen de l'apprendre par l'expérience. (Hertz, 1894, p. 67)⁵⁷

Ainsi, les images mentales au moyen desquelles nous prédisons et expliquons les phénomènes portent inévitablement la marque des esprits qui les engendrent.

On ne pourra éviter totalement les relations vides, car les images en comportent du seul fait qu'elles ne sont précisément que des images, et même des images de notre esprit particulier : elles doivent donc être déterminées aussi par les propriétés de *son mode de reproduction des objets*⁵⁸. (Hertz, 1894, p. 68)

Cependant, même si « nous ne pou[vons] éviter d'introduire dans nos images des relations auxiliaires inessentiels » (Hertz, 1894, p. 80), il est possible, selon Hertz, d'énoncer certaines exigences auxquelles nos images doivent se conformer au mieux. Il en énonce trois. Premièrement, elles doivent être *logiquement admissibles*, c'est-à-dire ne pas entrer en contradiction avec « les lois de notre pensée ». Deuxièmement, une image doit être *correcte*, c'est-à-dire refléter des relations fondamentales des objets extérieurs de manière à permettre de tirer des prédictions empiriques justes (sans quoi l'image ne correspondrait pas à la définition donnée ci-dessus). Enfin, une image doit être *distincte* et *appropriée*, c'est-à-dire à la fois « refléter plus de relations essentielles de l'objet » qu'une autre, et comporter « le plus petit nombre de relations superflues ou vides », donc être « la plus simple » possible. (Hertz, 1894, p. 68)

Or, si les questions de l'admissibilité logique et de la correction des images peuvent être tranchées objectivement,

quant à savoir si une image est appropriée, on ne peut pas, en général, trancher de façon univoque; au contraire, des divergences d'opinion peuvent exister. Une image peut offrir des avantages d'un certain point de vue, une autre d'un point de vue différent, et ce n'est que progressivement, grâce à la mise à l'épreuve de nombreuses images, que l'on pourra finalement, avec le temps, obtenir les images les plus appropriées. » (Hertz, 1894, p. 69)⁵⁹

⁵⁷ Boltzmann écrit ainsi : « Hertz a fait prendre conscience aux physiciens de quelque chose dont les philosophes ne doutaient plus depuis longtemps, à savoir qu'aucune théorie ne peut être objective, c'est-à-dire coïncider précisément avec la nature; bien plutôt, une théorie n'est qu'une image mentale des phénomènes [...] » (extrait de son essai de 1899, « Sur le développement des méthodes de la physique théorique dans la période récente », reproduit dans Boltzmann, 1974, pp. 90-91).

⁵⁸ Je souligne.

⁵⁹ De même, Poincaré reconnaît qu'« il peut se faire que [deux théories] expriment l'une et l'autre

Hertz reconnaît donc explicitement le caractère inévitablement idiosyncratique des images mentales au moyen desquelles nous représentons et prédisons les phénomènes. Non seulement elles portent la marque de l'esprit humain, mais encore elles sont relatives aux esprits individuels qui les produisent.

Les différentes présentations qui ont été proposées des principes de la mécanique en sont la preuve : elles expriment en effet la variété des images que les individus se forment des phénomènes mécaniques ; en retour, elles invitent à s'en former des images différentes :

les concepts fondamentaux de la mécanique, joints aux principes qui les lient, présentent l'image la plus simple que la physique puisse produire des objets du monde sensible et des processus qui s'y déroulent. Et comme *nous pouvons donner différentes présentations*⁶⁰ des principes de la mécanique en fonction des différents choix d'énoncés que nous prenons comme fondement, *nous obtenons plusieurs de ces images*⁶¹ des choses – images que nous pouvons tester et comparer entre elles sous le rapport de leur admissibilité, de leur correction et de leur appropriation. » (Hertz, 1894, p. 70)

De même que j'ai choisi, comme je l'annonçais à la section 2.2 du présent chapitre, de considérer les présentations explicites de la mécanique comme des expressions (au moins partielles) des versions de leurs auteurs, Hertz (1894) se propose, avant d'exposer sa propre version, d'étudier les deux grands types de présentations existantes (les « expositions scientifiques » des images)⁶² afin de clarifier les images qu'elles expriment, en les comparant du point de vue des trois exigences exposées :

Jusqu'ici, nous avons énuméré les exigences que nous posons à l'égard des images elles-mêmes ; tout autres sont les exigences que nous posons à l'égard d'une exposition scientifique de telles images. Nous exigeons de celle-ci qu'elle nous fasse prendre clairement conscience de la distinction entre les propriétés qui sont attribuées aux images au nom de l'admissibilité, celles qui le sont au nom de la correction, et celles qui le sont au nom de l'appropriation. Ce n'est que de cette façon que nous acquérons la possibilité d'apporter des modifications et des améliorations à nos images. (Hertz, 1894, p. 68)

des rapports vrais et qu'il n'y ait de contradiction que dans les images dont nous avons habillé la réalité. » (Poincaré, 1905, p. 75)

⁶⁰ Je souligne.

⁶¹ Je souligne.

⁶² Dans la mesure où Hertz ne distingue pas tant les présentations particulières que deux grands types de présentations, on peut bien les considérer comme deux *formulations* de la mécanique, ces dernières étant définies, si l'on veut, comme des classes d'équivalence de présentations. Cependant, comme on va le voir, les deux formulations dégagées par Hertz ne correspondent pas exactement aux formulations d'aujourd'hui : Lagrange y est rangé du côté de Newton.

Le projet de Hertz est donc, tout en prenant acte de l'impossibilité d'atteindre une représentation parfaitement objective, vraie et complète des phénomènes – sans quoi ce ne serait plus une *image* –, d'« exposer un agencement parfaitement déterminé des lois de la mécanique qui soit compatible avec l'état actuel de notre savoir [...] ». » (Hertz, 1894, p. 63)

4.2.2 Les deux images de la mécanique et la version de Hertz : le rejet du concept de force

Quelles sont les deux images qui se dégagent, selon Hertz, de l'histoire de la mécanique, et par quoi se différencient-elles ? La première image, la seule qui ait reçu un traitement systématique, est celle des présentations « ordinaires », fondées sur les concepts de force et de masse.

Une première image nous est fournie par la présentation ordinaire de la mécanique. Nous entendons par là celle que l'on trouve – sous des versions différentes dans le détail mais qui s'accordent sur l'essentiel – dans presque tous les manuels traitant de l'ensemble de la mécanique et dans presque tous les exposés qui couvrent l'ensemble du contenu de cette science. Cette présentation constitue la voie royale et la grande allée triomphale qui conduit la légion des écoliers vers les profondeurs de la mécanique ; elle suit exactement la marche du développement historique et de l'enchaînement des découvertes ; ses principales stations sont repérées par les noms d'Archimède, Galilée, Newton, Lagrange. Elle pose comme fondement, en tant que représentations données, les concepts d'espace, de force et de masse. La force y est introduite comme la cause du mouvement, qui lui préexiste et en est indépendante. (Hertz, 1894, p. 70)

La deuxième image, « d'origine beaucoup plus récente » (p. 82), ouvre la voie d'une « autre manière de penser » ; elle est fondée sur le concept d'énergie :

Influencée par l'immense impression que lui a causée la découverte du principe de conservation de l'énergie, elle aime à traiter les phénomènes qui sont de son ressort comme des conversions de l'énergie en de nouvelles formes, et à considérer comme son but ultime de rapporter les phénomènes aux lois de transformation de l'énergie. Cette façon de procéder peut aussi être appliquée dès le départ aux processus élémentaires du mouvement eux-mêmes ; par là même, une nouvelle présentation de la mécanique voit le jour, différente de la première, dans laquelle le concept de force s'efface d'emblée au profit du concept d'énergie. (Hertz, 1894, p. 82)

À la différence de la première, cette deuxième image n'a fait l'objet d'aucune présentation explicite et complète, et « il n'existe aucun manuel de mécanique qui se placerait dès le départ du point de vue de la théorie de l'énergie, et qui introduirait le concept d'énergie avant celui de force. » (Hertz, 1894, p. 82) Malgré cela, Hertz

affirme que l'on est bien en présence de deux images différentes, qu'un examen rationnel permet de départager des trois points de vue de l'admissibilité logique, de la correction et de l'appropriation. Sans entrer dans le détail de la critique que Hertz fait de ces deux images, je me contenterai ici d'en mentionner quelques aspects, particulièrement significatifs pour mon analyse.

Notons d'abord le point suivant : Hertz range sans hésitation Lagrange du côté de la mécanique de la force ; cela signifie que les progrès formels auxquels il a contribué ne changent pas, aux yeux de Hertz, l'image générale des processus mécaniques, qui consiste à les décrire et les expliquer au moyen du concept de force. Le « découpage » que Hertz opère entre les deux grandes manières de présenter la mécanique est donc guidé par une attention aux *concepts* qui donnent leur signification physique aux principes, ainsi qu'à l'ordre déductif des principes, et non à la forme des équations utilisées.⁶³

Or le concept de force est au cœur de sa critique de la première image ; en premier lieu, il reproche à toutes les présentations existantes de la mécanique de la force de ne jamais en donner une définition satisfaisante et claire. Il relève ainsi

la très grande difficulté que l'on éprouve à exposer en particulier l'introduction de la mécanique à des auditeurs perspicaces sans éprouver soi-même un certain embarras, sans avoir le sentiment de devoir s'excuser ici ou là, et sans ressentir le désir de survoler les débuts pour en arriver assez rapidement aux exemples qui parlent d'eux-mêmes. Je suppose que Newton lui-même a dû ressentir cet embarras lorsqu'il définit la masse de façon assez brutale comme le produit du volume par la densité. Je suppose que Thomson et Tait ont dû ressentir la même chose là où ils font remarquer qu'en réalité on a là une définition de la densité plus que de la masse, mais s'en contentent pourtant comme définition unique de la masse. Même Lagrange, je pense, doit avoir éprouvé cet embarras ainsi que le désir d'avancer à tout prix, lorsqu'il introduisit de but en blanc sa mécanique en expliquant qu'une force est une cause qui donne, « ou bien qui tend à donner », un mouvement à un corps ; il ne pouvait assurément pas ne pas sentir la difficulté logique d'une telle double détermination. (Hertz, 1894, p.73)

Le diagnostic de Hertz n'est pas sévère au point d'affirmer que la mécanique ordinaire est incohérente ou contradictoire et il admet que son « contenu » est logiquement ad-

⁶³ « Le concept de principe de la mécanique n'est [...] aucunement fixé de façon stricte. C'est pourquoi quand nous désignerons individuellement ces énoncés, nous conserverons certes leur appellation d'origine ; mais lorsque nous parlerons tout simplement des principes de la mécanique, nous ne voudrions pas que l'on entende par là ces quelques énoncés concrets, mais bien tout choix parmi ces énoncés et d'autres semblables – choix libre au demeurant – qui satisfasse la condition suivante : que l'on puisse développer à partir de là l'ensemble de la mécanique de façon purement déductive, sans aucun autre recours à l'expérience. » (Hertz, 1894, p. 70)

missible⁶⁴ ; cependant, un exposé de ce contenu logiquement admissible devrait pouvoir éviter ce genre de défaut.⁶⁵ Si la difficulté est si grande de donner une définition précise du concept de force, alors même que le système est logiquement admissible, il faut envisager, selon Hertz, de s'en passer : non seulement ce concept n'est pas indispensable à l'exposé de la mécanique, mais encore il est responsable de ce qu'elle a d'insatisfaisant dans sa forme. C'est donc d'abord au nom de la simplicité et de l'appropriation que Hertz propose de se passer du concept de force.

Le but de Hertz, dans sa propre présentation des principes, est de n'introduire dans cette image que des termes qui correspondent à des quantités observables. Les seuls concepts qu'il juge admissibles sont donc ceux de temps, d'espace et de masse ; énergie et force sont exclues du système. Il propose de remplacer les forces, censées être les causes du mouvement, par l'hypothèse de masses et de mouvements *cachés*. Il n'y aurait ainsi dans la nature rien que l'on puisse appeler « force » ou « énergie », mais simplement des masses et des mouvements, dont certains nous seraient imperceptibles. Les mouvements apparents s'expliquent ainsi tous par le même type de cause : il n'y a en fait que des systèmes à liaison, dont certaines sont cachées, mais aucun corps n'est soumis à l'action d'une quelconque force extérieure. Quand on croit qu'un corps obéit à une force, c'est qu'il est lié à d'autres corps, qui nous sont invisibles. Quand un point matériel s'écarte du mouvement rectiligne uniforme, c'est qu'il est lié à d'autres points invisibles, et non pas soumis à une force. Une telle explication est, pour Hertz, plus simple, plus naturelle et plus « intelligible » (Hertz, 1894, p. 95) : elle explique tous les phénomènes du mouvement par un seul type de cause.

Une première remarque s'impose. C'est, on vient de le voir, au nom de la simplicité et de l'intelligibilité qui en découle que Hertz cherche à se débarrasser du concept de force. On peut cependant légitimement affirmer, avec Mach, que l'idéal auquel obéit la préférence de Hertz pour des masses et des mouvements cachés sur l'hypothèse des forces, est des plus discutables : « les “forces” sont décidément préférables [...] aux “masses cachées” et aux “mouvements cachés”. Considérons un morceau de fer posé sur une table : les deux forces en équilibre, le poids du fer et l'élasticité de la table, sont très faciles à mettre en évidence. » (Mach, 1883, p. 319) On peut raisonnablement

⁶⁴ « nous sommes nous-mêmes convaincus que les lacunes existantes ne concernent que la forme, et que toutes les obscurités et incertitudes peuvent être évitées si on organise les définitions et les désignations de façon idoine, et si on s'exprime en outre avec circonspection. En ce sens, nous reconnaissons comme tout le monde l'admissibilité logique du contenu de la mécanique. » (Hertz, 1894, pp. 75-76)

⁶⁵ « Dans une science logiquement parfaite, en mathématique pure, une divergence d'opinion sur une question de ce genre est totalement inconcevable. » (Hertz, 1894, p. 74)

affirmer, avec Mach, que l'hypothèse des forces n'est pas moins simple et intelligible que celle de liaisons cachées.

Cependant, si Hertz voit dans la notion de force un défaut du point de vue de l'appropriation de l'image qu'elle sous-tend, ce n'est pas sur ce point qu'il affirme que la « troisième » image, la sienne, présente son plus grand avantage. Le corps des *Principes* tend à montrer qu'elle tire sa supériorité d'une plus grande correction, c'est-à-dire d'une meilleure adéquation empirique ; si l'image reposant sur le concept de force s'était montrée plus correcte (une fois le test de l'admissibilité logique passé), on aurait pu s'en satisfaire. La perfection de sa propre image provient, selon Hertz, du principe unique qu'il place au fondement du système, et qui est une généralisation de la loi d'inertie aux mouvements cachés :

si le système pouvait un instant rompre ses attaches, ses masses se disperseraient selon un mouvement rectiligne uniforme, mais [...], une telle dissociation étant impossible, elles restent du moins aussi près que possible de ce mouvement auquel elles aspirent. Cette loi fondamentale est dans notre image le premier énoncé empirique de la mécanique au sens propre ; elle en est aussi le dernier. Jointe aux hypothèses déjà admises des masses cachées et de la légalité des rapports, elle nous permet de dériver le reste du contenu de la mécanique de façon purement déductive. (Hertz, 1894, p. 98)

Comme on va le voir, nombreux ont été les savants à saluer dans l'œuvre de Hertz un accomplissement remarquable du point de vue de l'économie de la pensée et à reconnaître la beauté et l'économie de son système. Tous, cependant, s'accordent à affirmer que ce système, tout parfait soit-il, est pratiquement inutilisable et, en un certain sens, inintelligible.

4.2.3 Perfection formelle et intelligibilité de la mécanique

Les *Principes* de Hertz offrent en effet l'exemple remarquable d'une présentation où l'exigence de simplicité et la définition de l'intelligibilité qui l'accompagne entrent en conflit avec une autre forme d'intelligibilité de la théorie : non seulement il n'est pas évident que les « liaisons cachées » soient plus *faciles à concevoir* que les forces, mais encore, comme le dit Mach,

Lorsque l'on ne veut pas simplement accepter d'une façon générale l'hypothèse des masses et des mouvements cachés, mais que l'on cherche au contraire à l'utiliser pratiquement dans les problèmes *particuliers*, on doit, au moins dans l'état présent de nos connaissances physiques, en arriver à des fictions si extraordinaires et souvent si inimaginables, que l'emploi des accélérations *données* en devient de loin préférable. [...] Comme *programme idéal*, la mécanique de Hertz est plus belle et d'une plus grande unité que la mécanique ordinaire, mais celle-ci

l'emporte dans les applications [...]. (Mach, 1883, p. 255-256)⁶⁶

Notons, d'ailleurs, que cette appréciation de Mach s'accorde avec les buts affichés par Hertz :

Dans les détails, je n'ai en effet rien à exposer qui soit nouveau ou qui ne puisse être trouvé dans de nombreux livres. Ce qui, je l'espère, est nouveau, et la seule chose à laquelle j'attribue quelque valeur, est l'ordonnancement et l'agencement du tout, c'est-à-dire le côté logique, ou si l'on veut philosophique de la chose. Mon travail a atteint ou manqué son but selon que quelque chose a été ou non gagné dans cette direction. » (Hertz, 1894, p. 66).

Je voudrais, pour finir, souligner une véritable tension au sein même de l'entreprise de Hertz : d'une part, il défend une *Bildtheorie* et reconnaît le caractère idiosyncratique des théories, et, d'autre part, il cherche à proposer, au delà des présentations existantes, une méta-présentation débarrassée de tout effet de perspective⁶⁷, au point de la rendre inintelligible – au sens où elle n'a plus aucun ancrage dans l'expérience sensible – et inapplicable. Cela apparaît assez nettement dans la manière dont il revient, à la fin de l'introduction des *Principes*, sur l'idée de l'« appropriation » (voir ci-dessus, page 136) au nom de laquelle il a évalué les trois images :

nous n'avons jamais parlé d'appropriation que dans un sens particulier, à savoir au sens d'un esprit qui, sans égard pour la place fortuite de l'homme dans la nature, cherche à embrasser objectivement et à présenter de façon simple la totalité de notre connaissance physique ; mais [...] nous n'avons absolument pas parlé d'appropriation au sens de l'application pratique et des besoins humains. En ce qui les concerne, il n'est pas nécessaire que la présentation ordinaire de la mécanique, qui a été expressément imaginée pour eux, soit jamais remplacée par une présentation plus appropriée. *La présentation que nous introduisons ici entretient un peu le même rapport à la présentation ordinaire que celui qu'entretient la grammaire systématique d'une langue à une grammaire qui doit permettre aux débutants de se mettre au fait des nécessités de la vie quotidienne*⁶⁸ aussi vite que possible. On connaît l'étendue des différences entre les exigences qui pèsent sur l'une et l'autre ainsi qu'entre les agencements qui doivent être les

⁶⁶Duhem (1903b, 167) porte un jugement similaire, affirmant que la mécanique de Hertz, en raison même du fait que la mort lui a interdit de la mener à bien, est plus « le programme d'une doctrine » qu'une doctrine proprement dite. Il lui manque, pour être applicable et, de ce fait, susceptible de montrer sa fécondité ou sa stérilité, d'être « accompagnée de l'indication d'une méthode ». Voir (Mulligan, 1998) pour un examen de la réception des *Principes* de Hertz par ses contemporains.

⁶⁷C'est, de ce point de vue, l'ancêtre des approches formelles des théories que j'étudie au chapitre 5 : pourvus d'outils logiques et mathématiques dont Hertz ne disposait pas, les philosophes du Cercle de Vienne et leurs héritiers donnent comme critère de scientificité la possibilité, au moins en principe, de présenter les théories sous une forme axiomatique formelle rigoureuse.

⁶⁸Je souligne.

leurs, à l'une et à l'autre, si l'une comme l'autre doivent correspondre à leur but de façon aussi précise que possible. (Hertz, 1894, p. 112)⁶⁹

Comme on vient de le voir, le but que Hertz affirme poursuivre ici entre en conflit avec au moins deux autres vertus que l'on peut exiger de la présentation d'une théorie : la rendre intelligible et la rendre utilisable.⁷⁰

Concluons cet exposé en comparant rapidement la position de Hertz à celle de Mach : alors que ce dernier, en cherchant à fonder la signification de la mécanique sur les expériences sensibles des agents, affirme que ces différentes expériences – et les différentes manières de les formuler – sont fondamentalement équivalentes, Hertz établit une véritable distinction conceptuelle entre deux manières de présenter la mécanique, sans pour autant renoncer à l'idéal d'une présentation formellement parfaite. Ce sont donc deux conceptions très différentes des théories, aussi bien du point de vue normatif visant à caractériser la méthode et les buts que doit poursuivre l'entreprise scientifique que du point de vue descriptif de la nature même de la connaissance qui en résulte, qui donnent lieu à des versions ainsi qu'à des jugements méta-scientifiques et historiques différents. Je reviendrai sur les conclusions que l'on peut tirer de cette comparaison à la section 5, après avoir examiné deux autres versions.

4.3 Pierre Duhem (1903b), *L'évolution de la mécanique*

Pierre Duhem s'intéresse aux développements historiques de la mécanique énergétique⁷¹ et de la thermodynamique dans le cadre d'une entreprise de fondation de la physique entière sur une conception énergétiste du monde. Il développe par ailleurs (Duhem, 1914) une philosophie instrumentaliste, selon laquelle les théories ne doivent en aucun cas viser à fournir une explication des phénomènes en faisant l'hypothèse d'une quelconque réalité inobservable ; elles sont selon lui des représentations symboliques commodées résultant d'un double processus d'abstraction et de généralisation

⁶⁹La présentation de Hertz n'est pas non plus destinée à être un traité pédagogique ou un manuel à l'usage de ceux qui voudraient résoudre des problèmes de mécanique : « Pour diverses raisons, ce livre n'est pas, à vrai dire, tout à fait adapté comme première introduction ; mais c'est avec d'autant plus de conviction qu'il s'offre comme guide à qui connaît déjà, dans une certaine mesure, le contenu de la mécanique sous sa présentation ordinaire. Il espère pouvoir proposer à ce lecteur un point de vue qui offre au regard, en toute clarté, la signification physique, la parenté interne et la portée des principes de la mécanique ; un point de vue depuis lequel, en outre, le concept de force de même que les autres concepts fondamentaux de la mécanique apparaissent dépouillés de leurs derniers restes d'obscurité. » (Hertz, 1894, p. 64)

⁷⁰Je reviendrai sur cette tension au chapitre 4.

⁷¹Le terme « mécanique » dans le titre *L'évolution de la mécanique* désigne, comme le remarque Anastasios Brenner dans sa préface, la physique énergétique tout entière.

à partir des phénomènes.⁷² Dans ce qui suit, je me concentre sur quelques aspects de sa version de la mécanique lagrangienne telle qu'elle apparaît dans *L'évolution de la mécanique* (Duhem, 1903b) mais aussi dans sa critique de l'ouvrage de Mach (Duhem, 1903a). Comme il le souligne lui-même, et comme la lecture de leur œuvre respective ne tarde pas à le montrer, Duhem partage lui aussi avec Mach un grand nombre de convictions sur le but visé par une théorie scientifique, et notamment l'idéal d'économie⁷³, ainsi que le rejet des explications métaphysiques en science.⁷⁴

Cependant, la version de la mécanique de Duhem est en désaccord profond avec celle de Mach ; qui plus est, il ne s'accorde pas sur la manière de comprendre l'œuvre de Lagrange elle-même. Contre Mach, Duhem considère que la mécanique de la force, celle de Newton, mais aussi de Laplace et de Poisson, qui consiste à regarder « les corps comme formés de masses très petites isolées les unes des autres » n'est pas équivalente à la mécanique énergétique, qui traite « les corps comme des milieux continus dont les diverses parties se gênent les unes les autres en leurs divers mouvements et constituent des *liaisons* les unes pour les autres. » (Duhem, 1903a, p. 453) Mais, à la différence de Hertz, qui reconnaît aussi, on l'a vu, l'existence de plusieurs images différentes des phénomènes mécaniques, il ne range pas Lagrange du côté de la mécanique newtonienne.

Contrairement à Mach, qui « ne dissimule d'ailleurs pas ses préférences pour la méthode de Laplace et Poisson, qu'il nomme *méthode de Newton* », Duhem choisit résolument la voie énergétique. Ce qui définit selon lui cette méthode n'est pas en premier lieu de prendre le concept d'énergie comme fondamental, mais de représenter les corps et leurs mouvements d'une manière différente de la méthode newtonienne ; cela permet de se passer d'un certain nombre de suppositions qui rendent selon Duhem les explications fournies par la mécanique de la force aussi suspectes que celles des scolastiques faisant appel à des qualités et vertus occultes.⁷⁵

⁷² Pour une analyse de la philosophie des sciences de Duhem, voir (Brenner, 2000).

⁷³ « En tout domaine, le progrès de la Science a pour but de faire tenir le plus de réalité possible en une forme réduite le plus possible ; l'essence de ce progrès est une *économie* de plus en plus grande de la pensée. » (Duhem, 1903a, p. 445)

⁷⁴ On va voir dans ce qui suit que l'accusation de sombrer dans la métaphysique est une arme que différents « camps » utilisent, et qui n'a pas toujours le même sens.

⁷⁵ « Que l'on traite les attractions et les répulsions réciproques des points matériels comme des réalités irréductibles à la figure et au mouvement ; qu'on les considère, au contraire, comme les effets de mouvements, cachés encore à nos investigations ; il n'en reste pas moins que le physicien peut et doit invoquer dans ses raisonnements, non seulement des figures et des mouvements explicites, mais encore des *forces*, actuellement hétérogènes aux notions de la Géométrie et de la Cinématique. Par là, les mots : *Expliquer un phénomène physique* prennent un sens tout différent de celui que leur attribuaient les philosophes cartésiens ou atomistes ; l'explication qui s'arrête à la *force*, prise

Certes, la Physique dont Newton a tracé le plan et posé les bases, dont Bos-covitch a analysé la complète structure, est déjà admirable par la simplicité et l'ampleur de ses principes ; cependant, à côté de l'hypothèse fondamentale qu'il y a dans le monde : temps, étendue, masse et force, cette Physique n'admet-elle pas d'autres suppositions que l'on pourrait éliminer ? Au lieu de réduire la matière à un ensemble de points inétendus et isolés les uns des autres, ne pourrait-on y concevoir des corps étendus, variables de figure, capables de se toucher ? Au lieu de regarder toutes les forces comme des attractions et des répulsions réciproques, fonctions de la seule distance qui sépare les points qu'elles sollicitent, ne pourrait-on leur laisser une entière détermination, en accouplant seulement à chaque action une réaction égale directement opposée ? N'amènerait-on pas ainsi les principes de la Mécanique au plus haut degré de généralité qui se puisse concevoir ?

À cette construction de la Mécanique rationnelle, les plus grands géomètres du XVIII^e siècle contribuent ; Daniel Bernoulli, d'Alembert, Euler, pour ne citer que les plus illustres, attachent leur nom à quelque part de l'édifice ; mais à Lagrange on doit son complet achèvement. (Duhem, 1903b, p. 43)

Ce qui distingue la méthode de Bernoulli, d'Alembert, Euler et Lagrange – qui étaient rangés par Hertz, rappelons-le, dans la première image des phénomènes mécaniques – de celle de Newton n'est pas un remplacement du concept de force par un autre, ni un changement dans l'ordre des principes, mais l'introduction d'une nouvelle manière de représenter les systèmes physiques. Duhem note bien que « comme Leibniz, Lagrange regarde la notion de force comme une des notions premières de la Mécanique ; s'il invoque le mouvement, ce n'est pas pour expliquer la force, c'est seulement pour faire correspondre à cette idée, transcendante à la Géométrie, un symbole numérique capable de figurer dans les formules. » (Duhem, 1903b, p. 45). C'est le fait d'avoir « restitu[é] aux corps leur étendue, leur figure, la possibilité de glisser ou de rouler les uns sur les autres, voire de se déformer » (Duhem, 1903b, p. 45) qui conduit Lagrange, comme on l'a vu, à formuler le principe des vitesses virtuelles et le principe de d'Alembert en prenant en compte les liaisons des systèmes étudiés et à les représenter au moyen de coordonnées généralisées ; à partir de là, il « importe donc peu de connaître par le détail chacune des forces appliquées aux divers corps d'un système, son point d'application, sa grandeur, sa direction » (Duhem, 1903b, p. 49).

Poussons plus loin : Nous pouvons, si nous le voulons, ne considérer en Mécanique que des groupes de corps entièrement isolés dans l'espace ; il nous suffit, pour cela, de comprendre en un seul ensemble et le système particulier que nous

comme élément réellement ou provisoirement simple, a de l'analogie avec l'explication scolastique par les qualités et les vertus occultes. » (Duhem, 1903b, p. 42)

voulons étudier, et les corps dont l'influence sur ce système ne nous paraît pas négligeable. Alors nous n'aurons plus affaire qu'à des forces mutuelles s'exerçant entre les divers corps d'un même système ; ces forces intérieures sont supposées dépendre d'un potentiel, dont la connaissance rend inutile la connaissance des forces mêmes. Ainsi, la notion de *force*, après s'être fondue dans une notion plus ample, celle de *force généralisée*, perd, pour ainsi dire, son caractère premier et irréductible et apparaît comme une simple dérivation de la notion de *potentiel* ; telle est la conséquence naturelle des principes posés par Lagrange, conséquence qui s'accorde pleinement avec les vues profondes de Leibniz. (Duhem, 1903b, pp. 53-54)

Ainsi, Lagrange, selon Duhem, tout en acceptant le concept de force, propose une nouvelle manière de représenter les systèmes mécaniques, qui conduit progressivement à reléguer au second plan le concept de force, au profit des concepts énergétiques. C'est sur ce point que sa mécanique se distingue de celle de Newton, reprise par Laplace et Poisson.

Pour Duhem, on est en présence de deux théories ou manières générales de se représenter les corps et les phénomènes, non seulement pas équivalentes, mais même incompatibles. Après avoir cité Mach, à propos de l'équivalence des principes newtoniens et de ceux de la mécanique analytique, il affirme :

Nous ne souscrivons pas entièrement à ces jugements du professeur de Vienne ; nous ne croyons pas qu'il y ait toujours équivalence entre la méthode de Lagrange et la méthode que Laplace et Poisson ont tirée des principes de Newton ; nous pensons que l'extrême économie intellectuelle qui a présidé à la constitution de cette dernière méthode l'a trop appauvrie pour qu'elle puisse fournir une représentation satisfaisante de tous les phénomènes d'équilibre et de mouvement. (Duhem, 1903a, p. 455)

Cette différence de jugement entre Mach et Duhem à propos de l'équivalence ou non des méthodes newtonienne et lagrangienne repose sur le fait qu'ils ne tracent pas la frontière au même endroit. Plus précisément, Mach s'intéresse à la forme mathématique des principes et aux méthodes mises en œuvre, quand Duhem considère l'image du monde sous-jacente à ces méthodes. De ce point de vue, Lagrange passe, pour ainsi dire, dans l'autre camp. Alors que, pour Mach, la mécanique lagrangienne est une mécanique newtonienne, pour Duhem, la méthode analytique mise en œuvre par Lagrange en fait le père de la nouvelle mécanique, qui se distingue de la mécanique newtonienne en considérant les corps non plus comme composés de particules mais comme des milieux continus. Lagrange, dont on a vu précédemment qu'on pouvait dire qu'il avait confirmé les principes newtoniens en leur donnant des armes mathématiques très puissantes, sans rien changer à la signification de la théorie (aussi bien aux yeux de Lagrange lui-même, qu'à ceux de Mach et de Hertz), se trouve être,

dans la version que Duhem propose de l'histoire de la mécanique, le fer de lance des adversaires de la mécanique newtonienne.

La volonté d'unifier la physique sur la base des concepts énergétiques conduit ainsi Duhem à opérer un découpage nouveau des différentes représentations possibles des phénomènes du mouvement ; selon ce découpage, Lagrange ne se trouve pas du même côté que Newton. Pour définir le contenu d'une théorie, ce qui l'intéresse au premier chef n'est pas la forme mathématique des principes en tant que telle, mais la manière dont les corps eux-mêmes sont représentés et conçus. Ce faisant, il exprime une préférence nette en faveur des concepts énergétiques, que l'on peut raisonnablement critiquer comme une option métaphysique parmi d'autres possibles, consistant à affirmer la supériorité d'un standard d'intelligibilité sur un autre.

4.4 Une version contemporaine : Cornelius Lanczos (1970), *The Variational Principles of Mechanics*

Pour finir, et avant de dresser le bilan des versions examinées dans cette section, je propose de présenter brièvement celle qui ressort d'un des manuels de référence en mécanique aujourd'hui.⁷⁶ Le but de ce dernier examen est double. En premier lieu, le manuel de Lanczos étant une des sources de la connaissance en mécanique de beaucoup de physiciens théoriciens d'aujourd'hui, on peut le considérer comme l'expression d'une version de la mécanique assez largement partagée par les experts. En second lieu, la présentation de Lanczos me permet de renforcer encore le constat de la grande diversité des versions possibles, en particulier par le statut qu'elle accorde au principe de d'Alembert et par le « découpage » des formulations de la mécanique qui en résulte.

J'ai principalement opposé, dans ce qui précède, une mécanique de la force et une mécanique de l'énergie. Le principe de d'Alembert, comme on l'a vu, s'il permet à Lagrange de déduire des équations qui préfigurent les équations lagrangiennes d'aujourd'hui, est presque toujours rangé, dans les versions que j'ai présentées jusqu'ici, « du côté » de la mécanique newtonienne, dans la mesure où il ne fait pas intervenir de concept nouveau. En outre, on a vu que le maintien de ce principe au fondement de la mécanique analytique s'explique en partie, chez Lagrange, par le fait que ce dernier ne dispose pas du principe de moindre action sous une forme assez générale.

En raison même de l'inter-déductibilité des principes de la mécanique, le principe de d'Alembert n'est pas indispensable à la déduction des équations de la mécanique analytique : comme on l'a vu, il est possible d'obtenir les équations lagrangiennes

⁷⁶La première publication de l'ouvrage de Lanczos date de 1949. Il a été régulièrement réédité depuis.

du mouvement à partir d'un principe variationnel comme le principe de Hamilton.⁷⁷ Cependant, la plupart des manuels de mécanique qui procèdent à un examen approfondi de la structure de la mécanique, et cherchent à mettre en lumière l'articulation logique des principes, introduisent le principe de d'Alembert. L'ordre d'exposition du chapitre 2 du présent travail suit sur ce point celui du manuel de Lanczos (1970).⁷⁸ Ce manuel commence ainsi par présenter le principe de d'Alembert pour en déduire les équations du mouvement, ainsi que d'autres principes comme celui de la conservation de l'énergie, avant d'introduire les principes variationnels proprement dits.

Plusieurs raisons peuvent justifier un tel choix. Dans une perspective pédagogique visant à faciliter la compréhension, il est probable que la présentation du principe de d'Alembert comme une reformulation de la loi de Newton soit un premier pas plus facile à franchir avant de passer à la déduction des principes variationnels. Derrière ce que cette suggestion a de trivial, on peut voir l'idée plus profonde selon laquelle, quelle qu'en soit la raison (habitudes d'enseignement, aspect élémentaire des formalismes, ou encore aspect plus intuitif du concept de force), il est plus facile, pour comprendre la mécanique analytique, d'y entrer par la « fenêtre » newtonienne. En outre, mettre en évidence l'inter-déductibilité des différents principes permet d'enrichir notre compréhension des différents liens logiques entre les principes. C'est ce que suggère Jeremy Butterfield (2004), citant John Bell (1987, p. 77), à propos de l'enseignement de la relativité spéciale : « la route la plus longue donne parfois une plus grande familiarité avec le pays ».

Le principe de d'Alembert n'a cependant pas, dans la version de Lanczos, un statut simplement pédagogique. L'ouvrage de Lanczos est destiné à l'étude de la mécanique variationnelle, dont il souligne dès l'introduction les différences importantes qui la sépare de la mécanique vectorielle. Il fait ainsi remonter l'opposition entre les deux méthodes à celle entre Newton et Leibniz, le concept de *vis viva* de ce dernier

⁷⁷Certaines présentations de la mécanique, uniquement destinées à introduire les formalismes lagrangien ou hamiltonien en vue de l'étude d'autres théories qui en font usage, se passent souvent de l'exposé du principe de d'Alembert. C'est le cas, par exemple, du cours de Jean-Michel Raimond (2000), qui expose la formulation lagrangienne en guise d'introduction à la théorie de la relativité. Sans démontrer comment l'on peut, à partir du principe fondamental de Newton, obtenir les équations de Lagrange, il part du principe de Hamilton, en déduit les équations de Lagrange, dont il montre l'équivalence avec les équations newtoniennes pour des cas où elles sont toutes les deux applicables. En outre, dès l'introduction du cours, il affirme que le formalisme analytique n'ajoute rien de « conceptuellement nouveau ». On peut aussi déduire directement les équations de Lagrange à partir des principes newtoniens. Voir, par exemple, (Woodhouse, 1987, pp. 31-34, 41-47) et (Johns, 2005, chap. 2-2, 2-7). Dans (Desloge, 1982, pp. 522-523, 542-545, 554-557, 564-565), on passe progressivement de systèmes de moins en moins simples aux contraintes d'abord holonomes, puis non holonomes.

⁷⁸C'est aussi, à peu de choses près, l'ordre suivi par Goldstein (1950/2002).

étant considéré comme l'ancêtre de l'énergie. Grand admirateur de la méthode variationnelle, Lanczos insiste, dans son introduction, sur le caractère remarquable de l'équivalence des deux méthodes et de la possibilité même de représenter un phénomène « dirigé » comme le mouvement au moyen de quantités scalaires (Lanczos, 1970, p. XXI). Son but est ainsi de présenter la méthode variationnelle, fondée sur un principe de minimum et sur le concept d'énergie. Autrement dit, l'ouvrage entier se présente comme une version de la mécanique variationnelle, et non de la mécanique newtonienne, qu'on ne retrouve qu'à titre de corollaire.

Or, bien que le principe de d'Alembert ne soit pas, en tant que tel, un principe variationnel, et malgré le fait qu'il soit initialement formulé au moyen du concept de force et non de celui d'énergie, Lanczos lui accorde un statut fondamental pour la mécanique variationnelle.

L'importance de l'équation $[\mathbf{F} + \mathbf{I} = 0]$ réside dans le fait qu'elle est *plus* qu'une reformulation de l'équation de Newton. Elle est l'expression d'un *principe*. Nous savons que l'annulation d'une force en mécanique newtonienne signifie l'équilibre. Donc l'équation $[\mathbf{F} + \mathbf{I} = 0]$ dit que l'addition de la force d'inertie aux autres forces agissantes produit l'équilibre. Mais cela signifie que si nous avons un critère pour l'équilibre d'un système mécanique, nous pouvons immédiatement étendre ce critère à un système qui est en mouvement. Tout ce que nous avons à faire est d'ajouter la nouvelle « force d'inertie » aux autres forces. Par cette méthode *la dynamique est réduite à la statique*. (Lanczos, 1970, p. 89)

Cette signification supplémentaire qu'apporte le principe de d'Alembert à la mécanique, Lanczos affirme qu'elle est déjà contenue dans la restriction qui se trouve au fondement de l'énoncé du principe du travail virtuel et du principe de d'Alembert (voir ci-dessus, chapitre 2, page 54). Comme je l'ai signalé (chapitre 2, note 35 page 55), Lanczos appelle cette restriction le « postulat *A* » :

Nous pouvons demander quelle est la signification *physique* du principe de d'Alembert. De la définition de la « force effective » par l'équation $[\mathbf{F}_k^e = \mathbf{F}_k + \mathbf{I}_k]$ il suit que cette force est nulle dans le cas d'une particule libre, et égale à la force négative de réaction si la particule est sujette à des contraintes. Donc l'application du principe du travail virtuel aux forces effectives \mathbf{F}^e est équivalente à l'hypothèse selon laquelle « le travail virtuel des forces de réaction est nul pour tout déplacement virtuel en harmonie avec les contraintes ». Nous sommes donc renvoyés au même « Postulat *A* » que nous avons rencontré précédemment [...]. Le principe de d'Alembert généralise ce postulat du champ de la statique au champ de la dynamique, sans aucune altération. (Lanczos, 1970, p. 91)

En élevant cette restriction au rang de postulat, Lanczos affirme que tous les principes de la mécanique variationnelle, en tant qu'ils ne sont que des reformulations du principe de d'Alembert, en découlent :

Ce postulat n'est pas restreint au royaume de la statique. Il s'applique également à la dynamique, quand le principe du travail virtuel est convenablement généralisé au moyen du principe de d'Alembert. Puisque tous les principes variationnels fondamentaux de la mécanique, les principes d'Euler, Lagrange, Jacobi, Hamilton, ne sont que des *formulations mathématiques diverses du principe de d'Alembert*⁷⁹, le Postulat *A* est en fait le *seul*⁸⁰ postulat de la mécanique analytique, et est par conséquent d'une importance fondamentale. [Lanczos ajoute en note : Ces scientifiques qui prétendent que la mécanique analytique n'est rien d'autre qu'une formulation mathématique différente des lois de Newton doivent supposer que le Postulat *A* est déductible des lois du mouvement de Newton. L'auteur ne voit pas comment cela peut être fait. La troisième loi du mouvement, « l'action égale la réaction », n'est certainement pas assez large pour remplacer le Postulat *A*.] » (Lanczos, 1970, pp. 76-77)

Ainsi, Lanczos affirme bien que la mécanique variationnelle *dit* quelque chose de plus que la mécanique newtonienne, que son contenu est modifié par l'introduction des principes variationnels, mais il considère que la vraie nouveauté réside dans le postulat sur lequel repose le principe de d'Alembert et ses reformulations variationnelles ; tout ce que dit la mécanique variationnelle en plus de la mécanique newtonienne est contenu dans ce postulat.

On a vu dans ce qui précède plusieurs manières d'apprécier le travail de Lagrange et de comprendre la mécanique analytique. Dans presque tous les cas, le principe de d'Alembert est considéré comme une reformulation du principe de Newton, sans signification nouvelle. Lanczos fait du principe de d'Alembert *le* principe fondamental de la mécanique variationnelle, dont l'introduction résulte en un changement conceptuel au sein de la mécanique. Pour lui, si les principes variationnels sont conceptuellement différents de ceux de la mécanique vectorielle, ce n'est pas seulement en raison de l'introduction du concept d'énergie ni de la modification de l'architecture globale de la mécanique, mais parce qu'ils sont la conséquence d'une affirmation nouvelle – le postulat *A* –, que d'autres, comme Butterfield (2004)⁸¹, traitent comme une restriction.

Après ce tour d'horizon de quelques versions possibles de la mécanique classique, je propose, à la section suivante, d'en dresser le bilan. J'en examinerai quelques conséquences pour l'analyse de la relation entre changement formel et changement conceptuel.

⁷⁹ Je souligne.

⁸⁰ Italiques d'origine.

⁸¹ Voir chapitre 2, page 54.

5 Bilan de l'analyse des versions : différences conceptuelles et différences formelles

J'ai fait l'hypothèse, dans ce chapitre, que les utilisateurs d'une théorie se forment, au cours de leur apprentissage et de leur pratique, un ensemble de représentations mentales au moyen desquelles ils développent et appliquent cette théorie pour prédire et expliquer les phénomènes. Cette hypothèse, destinée à rendre compte de l'activité théorique, s'accompagne de la conviction selon laquelle les théories, en tant qu'elles sont des outils de représentation et d'inférence, ne sont pas indépendantes des raisonnements des agents et ne fonctionnent qu'en tant qu'elles sont *comprises*.

Cela implique que le contenu même d'une théorie ainsi conçue n'est pas identifiable indépendamment des versions de ses utilisateurs. Étudier différentes présentations de la mécanique comme les expressions d'autant de versions m'a ainsi permis de mettre en évidence deux choses, sur lesquelles je souhaite revenir dans cette dernière section. En premier lieu, la grande diversité de ces versions permet de montrer que, selon la manière dont elle est présentée, la mécanique n'a pas exactement la même signification. Autrement dit, j'ai montré qu'il existe des différences conceptuelles entre les versions de la mécanique – et les présentations qui les expriment. En second lieu, certains jugements explicites exprimés par les auteurs étudiés permettent de mettre en évidence que l'identification et l'évaluation de ces différences conceptuelles dépendent elles-mêmes de la manière dont chacun utilise et comprend la théorie et ses différentes formulations.

Dans un premier temps (section 5.1), je reviens sur le statut des différents facteurs dont j'ai supposé qu'ils déterminent, au moins partiellement, les versions des agents. Je propose ensuite (section 5.2) un aperçu récapitulatif des principales différences identifiées entre les versions examinées. Cela va me permettre, enfin (section 5.3), de revenir sur la question de la distinction entre une différence conceptuelle (ou un changement conceptuel) et une différence formelle (ou un changement formel).

5.1 Idéal scientifique et « psychologie » des agents

J'ai supposé, dans la définition que j'ai proposée de la notion de version à la section 2 du présent chapitre, que la version de chaque individu est au moins partiellement déterminée par un ensemble de facteurs que l'on peut appeler « contextuels » et « psychologiques » : l'apprentissage qu'il a suivi, ses connaissances d'arrière-plan, mais également ses talents personnels, ses habitudes de raisonnement, ses engagements théoriques et ses préférences philosophiques. J'ai suggéré que, parmi les experts eux-mêmes, on peut distinguer différents « types d'esprits » qui, tout en respectant

les relations logiques entre les concepts de la mécanique, ne les utilisent et ne les comprennent pas de la même manière.

L'analyse de la section 4 m'a permis de renforcer cette hypothèse, en montrant que les versions de la mécanique exprimées par différents savants de la fin du dix-neuvième et du vingtième siècles semblent elles-mêmes partiellement déterminées par des facteurs de ce type, auxquels nous avons un accès plus ou moins direct. Parmi eux, certains sont des choix philosophiques ou scientifiques explicites, que l'on peut appeler, en empruntant cette expression à Amy Dahan-Dalmedico (1990), l'« idéal scientifique » de l'auteur ; d'autres relèvent de déterminants psychologiques dont le rôle est plus difficile à cerner, mais dont la prise en compte semble utile pour éclairer les divergences d'opinion.

L'idéal scientifique de chaque auteur se caractérise par un ensemble d'exigences et de convictions philosophiques à propos de la nature et des buts d'une théorie scientifique. Dans presque toutes les présentations que j'ai examinées, on a vu qu'un certain idéal de simplicité, d'économie, ou encore de généralité, est souvent invoqué pour justifier diverses appréciations sur les différents principes de la mécanique, sur les différentes manières de les ordonner, et sur la contribution respective des acteurs de son développement historique. Cet idéal de simplicité, ainsi qu'un certain « esprit anti-métaphysique », pour reprendre les termes de Mach, quoique revendiqués par tous, peuvent cependant être compris diversement selon les cas, et donner lieu à des conceptions divergentes de la signification physique et de l'importance relative des différents principes. C'est, par exemple, au nom de cet esprit anti-métaphysique que Mach défend le concept de force, en raison même de son caractère intuitif dû à son ancrage dans une expérience sensible, et qu'il s'oppose à l'idée selon laquelle les principes variationnels apportent une signification nouvelle à la théorie. C'est pourtant au nom du même type d'exigences que Hertz cherche à débarrasser la mécanique du concept de force, et que Duhem voit dans la représentation des corps qui conduit naturellement à l'adoption des concepts énergétiques un accomplissement remarquable.

Le consensus autour des valeurs mentionnées est donc principalement verbal. Notons que l'on aurait du mal à imaginer qu'un scientifique puisse refuser l'idée selon laquelle la science doit viser la simplicité et éviter les hypothèses dont l'ancrage empirique est douteux et qui ont une saveur métaphysique. Cependant, le fait que les différentes manières de comprendre ces valeurs les font parfois entrer en conflit suggère que l'idéal scientifique de chacun est en fait partiellement déterminé par d'autres facteurs, d'ordre psychologique et contextuel, qui font rarement l'objet d'un examen explicite. Hertz, tout comme Mach, reconnaît d'ailleurs l'importance de la psychologie des savants dans leur travail scientifique. Cette « psychologie » est un objet bien difficile à saisir pour l'historien, qui ne peut qu'émettre des hypothèses

invérifiables sur la façon dont la formation et la spécialité scientifiques de chacun, ses talents propres, ainsi que des éléments d'un contexte scientifique et linguistique plus ou moins large (allant de la société à la communauté scientifique ou « école »), déterminent ses options philosophiques, ainsi que le type de raisonnement qu'il mène avec le plus d'aisance.

Tout en reconnaissant le caractère conjectural des explications en appelant à ces éléments contextuels, le philosophe peut cependant affirmer qu'ils jouent un rôle, certes difficile à cerner, dans la vie cognitive et dans l'activité théorique de chacun. Comme je l'ai souligné au début de mon exposé de la notion de version (page 93), et comme cela doit être apparu dans l'examen que je viens de proposer, supposer comme je le fais que la version de chacun est partiellement déterminée par cet ensemble de facteurs que j'appelle « psychologiques » ne veut pas dire que ce qui m'intéresse au premier chef est la relation entre ces facteurs et la version des agents. Comme je l'ai dit à la section 2.2, le seul accès que nous ayons à ces versions est la manière dont elles se manifestent dans la pratique des agents (leur utilisation des équations de la mécanique pour résoudre des problèmes) et, le cas échéant, dans la présentation explicite qu'ils font de cette théorie. C'est aux versions ainsi exprimées, et à la manière dont elles invitent à reconsidérer le contenu de la mécanique, que je m'intéresse. Cependant, décrire les versions comme des ensembles de représentations mentales implique de les considérer comme des éléments de la vie psychologique des agents.

Cela permet d'expliquer, au moins partiellement, la très grande diversité des positions (explicites) adoptées par les auteurs que j'ai étudiés. Je me propose à présent de dresser un bilan synthétique des différentes options choisies par chacun d'entre eux. Cela va me permettre de montrer la pertinence d'un examen des versions de la mécanique pour une analyse du contenu de cette théorie (et de ses formulations).

5.2 Bilan récapitulatif des versions étudiées

Les présentations de la mécanique que j'ai examinées se distinguent les unes des autres par plusieurs aspects. Les différentes options choisies ne sont pas indépendantes ; toutefois, on l'a vu, leur assortiment varie, de telle sorte que, si l'on peut grossièrement opposer deux grands types de versions, ce ne sont pas des « blocs » nettement délimités.

5.2.1 Architecture des principes et ordre des concepts

En premier lieu, les présentations se différencient par les deux aspects principaux qui permettent de caractériser une version de la mécanique : l'architecture des principes et l'ordre des concepts. Si ces deux aspects sont liés, ils se « combinent »

différemment selon les cas.

Architecture des principes On présente la plupart du temps la mécanique sous la forme d'un ensemble d'énoncés organisés de manière déductive. Les présentations se distinguent les unes des autres par les énoncés qu'elles instituent en principes fondamentaux, par l'ordre dans lequel elles déduisent les autres, et par le statut qu'elles leur donnent (postulat, théorème, etc.). De ce point de vue, deux grands types de présentations s'opposent, celles qui conservent les principes newtoniens comme principes premiers, et celles qui voient dans le principe de Hamilton, ou plus généralement dans un principe variationnel, le véritable fondement de l'édifice. Certains auteurs considèrent aussi que l'authentique principe fondamental est celui de d'Alembert.

Ordre des concepts On peut aussi distinguer deux grands types de versions sur la base des concepts dont la théorie est supposée tirer sa signification. Certains auteurs considèrent ainsi que le concept d'énergie tire sa signification du concept de force, dont il est dérivé, d'autres prennent au contraire le concept d'énergie comme premier. Hertz, quant à lui, propose de se passer des deux. Le choix des concepts premiers va souvent de pair avec celui des principes fondamentaux : placer le principe de Hamilton au fondement de la dynamique, c'est généralement admettre que l'énergie en est le concept premier. Ce n'est cependant pas toujours le cas : on peut placer, comme Mach ou Lagrange, le principe de d'Alembert, ou même le principe de Hamilton, au fondement de la mécanique pour des raisons d'économie formelle, sans pour autant considérer qu'il soit premier du point de vue de la signification, et, au contraire, affirmer qu'il tire sa signification de celui de Newton, dont il est alors considéré comme une réécriture.

5.2.2 Jugements méta-scientifiques et historiques

Les auteurs étudiés dans ce chapitre assortissent généralement leur présentation des principes et des concepts de la mécanique d'un ensemble de jugements explicites sur la relation (équivalence ou différence) entre les principes et, le cas échéant, entre les différentes présentations de la mécanique (dont la frontière n'est pas toujours située au même endroit). Enfin, ces jugements sont souvent eux-mêmes liés à une certaine appréciation des grandes étapes de l'histoire de la mécanique.

Jugement porté sur la relation entre les principes Certaines présentations insistent sur l'équivalence des différents principes – ou plutôt des différentes formes que prennent les principes fondamentaux – quand d'autres affirment que le principe

de d'Alembert, ou encore celui de Hamilton, *dit* quelque chose de plus que celui de Newton. Ce jugement sur l'équivalence ou non des formulations est souvent lié à certains choix relatifs à l'architecture des principes : conserver les principes newtoniens au fondement de la mécanique revient à affirmer que tous les développements ultérieurs sont déjà contenus dans ces principes. Inversement, placer le principe de Hamilton au fondement de la théorie implique souvent l'affirmation selon laquelle ce principe dit plus de choses que les principes newtoniens. Cependant, il convient de faire la même remarque que précédemment (voir ci-dessus le paragraphe sur l'ordre des concepts) : on peut considérer qu'un principe variationnel est plus économique formellement sans affirmer qu'il est autre chose qu'une réécriture des lois de Newton.

Tracé des frontières entre les différentes présentations et évaluation des versions qu'elles expriment

Certains auteurs prennent acte de la possibilité de présenter la mécanique de différentes manières, et en font l'objet de réflexions explicites. Leur version s'exprime alors aussi par la manière dont ils présentent et comprennent les autres versions exprimées au cours de l'histoire, ou par leurs contemporains. La plupart du temps, ils opposent deux grands types de présentations ; mais le tracé des frontières varie de version à version (puisque les distinctions mentionnées ci-dessus ne se recoupent pas toujours). Si l'on prête attention aux concepts fondamentaux, on opposera, avec Hertz, une mécanique de la force et une mécanique de l'énergie. Si l'on prend comme critère principal le langage mathématique dans lequel sont exprimés les principes fondamentaux, à partir desquels on obtient les équations décrivant le mouvement des systèmes, on distinguera entre une mécanique des équations différentielles et une mécanique variationnelle. On peut aussi, avec Duhem, opposer deux manières de voir les corps (comme composés de corpuscules ou comme des solides qui ne peuvent s'interpénétrer), ce qui ne recoupe pas exactement la distinction entre force et énergie. Enfin, le découpage standard des manuels d'aujourd'hui, que la présentation du chapitre 2 a suivi, oppose la mécanique newtonienne ou vectorielle à la mécanique analytique ou variationnelle, sans qu'une distinction soit opérée entre ces deux derniers qualificatifs. Le critère qui préside à un tel découpage est alors la forme des équations utilisées pour représenter le mouvement des systèmes.

Appréciation des grandes étapes de l'histoire de la mécanique, et en particulier de la contribution de Lagrange

Les différents « découpages » se recouvrent si peu que, selon les cas, Lagrange se trouve d'un côté ou de l'autre de la frontière dressée entre deux grands types de présentations. Pour Hertz, il est rangé dans le camp de Newton, celui d'une mécanique de la force ; Duhem, accordant moins d'importance aux concepts qu'à la manière de représenter les systèmes physiques,

admet que la mécanique lagrangienne est une mécanique de la force, mais fait de Lagrange le père de la mécanique énergétique, fondée sur une nouvelle manière de représenter les corps. L'usage d'aujourd'hui, qui consiste à appeler « formulation lagrangienne » un type de présentation fondée sur le principe de Hamilton, peut être interprété comme l'expression d'un certain consensus qui consiste à attribuer à Lagrange la véritable paternité de la mécanique variationnelle.

En conséquence, selon le « découpage » adopté et selon le « camp » dans lequel Lagrange est rangé, on jugera tantôt que ses travaux consistent en un développement formel de la mécanique, et tantôt qu'il a opéré un véritable bouleversement conceptuel. C'est sur ce point que je me propose de revenir pour clore ce chapitre.

5.3 Changement formel et changement conceptuel

Une des questions que l'on retrouve sous la plume de tous les auteurs étudiés dans ce chapitre est celle de la *signification physique* de la théorie et de ses différents principes. C'est elle qui décide de l'équivalence ou de la différence des principes, et qui nous dit quel concept est premier et donne sa signification à l'ensemble de la théorie – ou à l'une de ses formulations, dans le cas où l'on juge qu'elles ne sont pas équivalentes conceptuellement. L'idée partagée par tous est qu'un changement conceptuel consiste en une modification de la signification – ou du contenu – physique de la théorie. Cependant, alors qu'il peut sembler évident, avant toute analyse, que le contenu physique d'une théorie doit pouvoir être identifié à l'ensemble de ses conséquences empiriques, on vient de voir que la question est beaucoup plus délicate. L'ensemble des choix relatifs à chacun des points récapitulés dans la section précédente peut être considéré comme l'expression d'une certaine conception de *la distinction entre une avancée conceptuelle et un développement formel*.

Reprenons les différentes opinions exprimées à propos de la contribution de Lagrange. Lagrange lui-même, on l'a vu, considère qu'il n'a pas modifié la signification de la mécanique. Chez lui, ce jugement s'accompagne de la conviction que cette signification est tout entière contenue dans les principes de Newton, et que les modifications ultérieures ne peuvent être que formelles. Pour autant, le changement introduit par Hamilton est souvent considéré (par exemple par Pulte, voir ci-dessus, page 119) comme un changement conceptuel. Selon cette opinion, un changement conceptuel est identifiable à un changement dans l'ordre déductif des principes ; en revanche, un développement du langage mathématique dans lequel les principes et les équations sont exprimés ne compte pas, dans cette perspective, comme un changement conceptuel s'il ne débouche pas sur une telle réorganisation.

La diversité des versions exprimées par différents savants, tous post-hamiltoniens,

m'a permis de montrer qu'une telle conception du changement conceptuel est loin de faire l'unanimité. On peut, comme Mach, admettre que le principe de Hamilton est la formule la plus générale de la mécanique sans pour autant considérer que la réorganisation qui en résulte change la signification physique du système. De même que l'on pouvait, avant de posséder les outils mathématiques nécessaires, chercher, comme Maupertuis, à fonder la mécanique sur un principe de minimum, on peut être newtonien après Hamilton et considérer qu'aucun changement conceptuel n'est intervenu dans l'histoire de la mécanique.

Inversement, on peut aussi affirmer que c'est la contribution de Lagrange, et non celle de Hamilton, qui constitue l'avancée conceptuelle la plus importante. Un tel jugement repose sur l'idée selon laquelle la reformulation analytique des équations de la mécanique, parce qu'elle ouvre la voie à une nouvelle manière de représenter, de prédire et d'expliquer les phénomènes du mouvement, constitue en elle-même une évolution conceptuelle. Selon ce dernier point de vue, il n'est pas illégitime d'attribuer à Lagrange la paternité de la formulation lagrangienne d'aujourd'hui.

La difficulté de trancher entre ces différents jugements est en partie liée à l'absence, soulignée par Anouk Barberousse (2008), de toute distinction claire, en physique théorique, entre la part des mathématiques et la part de la physique. Si l'on considère que les équations différentielles sont des représentations des phénomènes, qu'elles disent quelque chose à propos du monde, on sera tenté de voir dans une nouvelle forme d'équation une nouveauté conceptuelle, et dans l'enquête mathématique permettant le développement formel de la mécanique une exploration de sa signification physique. Dans une telle perspective, un changement formel (mathématique) peut bel et bien modifier la signification *physique* d'une théorie.

On peut finalement tirer de ces considérations une remarque sur la relation entre histoire et philosophie des sciences : la tâche de l'historien visant à retracer les étapes de la naissance de la mécanique analytique et celle du philosophe cherchant à en clarifier la structure conceptuelle sont ici profondément liées et pratiquement indissociables. Selon que l'on accorde plus d'importance à l'évolution des formalismes mathématiques, à l'architectonique des principes, à la définition des concepts, ou encore aux options métaphysiques sous-jacentes aux différentes présentations de la mécanique, la chronologie change d'aspect. Réciproquement, selon la manière dont on décrit la contribution des différents acteurs de l'histoire de la mécanique (de la réécriture formelle à l'innovation conceptuelle), l'évaluation des différences entre mécanique newtonienne et mécanique analytique, et la définition même de cette dernière, changent. En raison même de la difficulté d'établir une distinction nette entre forme et contenu d'une théorie, le changement théorique lui-même n'est pas facile à dater.

C'est d'ailleurs la remise en cause de l'opposition classique entre forme et contenu

empirique, sur laquelle se fonde généralement l'appréciation historique du changement théorique, qui conduit le philosophe Ian Hacking (1982, 1983a, 1992a,b, 2003) à trouver dans la notion de « style de raisonnement » proposée par l'historien Alistair Crombie (1994)⁸² une catégorie éclairante pour l'histoire des sciences.⁸³ Elle permet en particulier, selon lui, de rendre compte de la manière dont l'invention d'un langage mathématique, sans forcément s'accompagner d'une nouveauté empirique, introduit une nouvelle manière de raisonner et de représenter les phénomènes ; quand bien même les résultats que ces raisonnements permettent d'atteindre sont en principe déductibles de ce que l'on connaissait déjà, la prise en compte de l'aspect essentiellement créatif de l'invention théorique interdit selon Hacking de prendre « la déductibilité comme critère d'équivalence » :

Le fait d'imaginer les déductions ne laisse pas les choses identiques. Weinberg⁸⁴ imagine un tableau analogue à celui d'un écolier qui doit résoudre un exercice de géométrie euclidienne. Si l'enfant résout le problème, il écrit Q.E.D. à la fin de sa démonstration, *quod erat demonstrandum*. Dans une science qui se développe, le « *quod* » n'est généralement pas là avant la démonstration. Les grandes figures de ce qui a un jour été appelé la mécanique rationnelle, des hommes comme Laplace et Lagrange travaillant aux alentours de 1800, étaient en un sens en train de tirer les conséquences des lois du mouvement et de la gravitation de Newton. Mais ils devaient inventer les mathématiques susceptibles de le faire. Ils devaient inventer le langage dans lequel les conclusions pourraient être exprimées. Ils devaient articuler la théorie. Ils n'étaient pas seulement en train de dessiner la liaison entre des points pour parachever une image. Ils devaient placer les points. Je ne fais ici qu'évoquer des questions incroyablement difficiles. Déductibilité, traductibilité et équivalence ne sont pas des idées transparentes. (Hacking, 1999, p. 109)⁸⁵

Le « style » de Lagrange peut ainsi bel et bien être considéré comme l'introduction d'une nouvelle manière de raisonner et de comprendre la mécanique : en proposant

⁸²Crombie parle de « styles de pensée scientifique ».

⁸³Pour une évaluation de l'utilité de la catégorie de style en histoire des sciences, voir l'article de Jean Gayon (1996b), qui note en particulier qu'elle permet de rendre compte de la liaison entre la « singularité d'une forme particulière » et l'« universalité nécessaire à la notion d'objectivité » et de rendre justice, en conséquence, au « processus historique complexe par lequel des connaissances étroitement dépendantes d'un contexte viennent à s'universaliser. »

⁸⁴Hacking fait référence aux propos du prix Nobel de physique Steven Weinberg (dans le cadre de l'affaire Sokal, née du canular de ce dernier, Sokal, 1996) qui affirme que « si jamais nous découvrions des créatures intelligentes sur quelque planète lointaine et traduisions leurs travaux scientifiques, nous découvririons qu'eux et nous avons découvert les mêmes lois. » (Weinberg, 1996, p. 14, cité par Hacking, 1999, p. 107).

⁸⁵La traduction citée est celle de Baudouin Jurdant, dans l'édition indiquée dans la bibliographie.

une nouvelle *forme de représentation*, il ouvre la voie à des *types d'inférences* nouveaux.⁸⁶

Au terme du chapitre 2 (page 82), la remise en cause de la pertinence de la notion classique de théorie m'a permis d'envisager une nouvelle cartographie de l'activité théorique, centrée sur les types de représentation utilisés par les agents, et sur les raisonnements et pratiques qui leur sont associés. De façon analogue, dans le présent chapitre, la remise en cause de la distinction entre changement formel et changement conceptuel invite à proposer de nouvelles catégories d'analyse en histoire des sciences, qui remettent en cause les chronologies habituelles.

Conclusion

Le chapitre 2 m'a conduit à montrer que l'identification du contenu d'une théorie à la clôture déductive d'un ensemble d'hypothèses ne permet pas de rendre compte de son utilisation comme outil de représentation et d'inférence. Le présent chapitre, en adoptant la perspective des agents, a proposé une analyse des différences conceptuelles qui existent au sein de la mécanique, d'un double point de vue diachronique et synchronique.

L'hypothèse des versions est sous-tendue par la conviction que le contenu d'un ensemble d'hypothèses théoriques ne peut pas être identifié sans la prise en compte de la manière dont les agents les utilisent pour représenter et pour tirer des inférences à propos des phénomènes. Par la mise en œuvre de cette hypothèse, j'espère avoir montré que, dès lors que l'on prête attention à ce que j'ai identifié, à la section 1, comme des différences conceptuelles entre les formulations d'une « même » théorie, on gagne à franchir un pas supplémentaire, et à prendre en compte également certaines différences *entre les agents*. La motivation initiale de l'examen des formulations de la mécanique était de remettre en cause la définition logico-empirique du contenu des théories et de montrer l'importance de la forme sous laquelle elles sont présentées *pour les raisonnements des agents*. Or, l'examen des versions de la mécanique que je viens de proposer a mis en évidence que la mécanique, sous ses différentes formulations, n'est pas comprise et utilisée de la même manière par différents experts. Autrement dit, il m'a permis de montrer que les différences de formulation n'ont pas exactement les mêmes conséquences pour les raisonnements de différents agents. Si l'on renonce à considérer que le contenu d'une théorie est défini par l'ensemble de ses conséquences déductives en principe, et que l'on s'intéresse aux inférences que les

⁸⁶Le lien entre forme de représentation et type d'inférences est au cœur de l'analyse du chapitre 7 du présent travail.

agents peuvent effectivement tirer au moyen de ses différentes formulations, on doit donc aussi admettre que ce contenu n'est pas exactement le même pour des agents différents.

Encore une fois, insistons sur la chose suivante : je m'intéresse, dans ce travail, aux conséquences de la prise en compte des utilisateurs des théories pour une analyse de leur contenu. Si deux agents, ou deux groupes d'agents, n'utilisent pas les représentations fournies par une théorie de la même manière, cela implique que, dans une certaine mesure, le contenu de cette théorie n'est pas le même pour ces deux (groupes) d'agents. J'introduis, à titre d'hypothèse, l'idée selon laquelle l'utilisation d'une théorie par les agents est le symptôme – et l'origine – d'une certaine organisation de ce que je propose de considérer comme un réseau de représentations mentales (leur version). En tant que telles, ces versions font partie de leur vie psychologique et sont, donc, déterminées par certaines particularités psychologiques des agents. Cependant, la façon dont ces particularités psychologiques déterminent la version des individus ne m'intéresse pas directement. Pour le dire autrement, on peut supposer que deux agents, pour des raisons psychologiques tout à fait différentes, ont la même version de la mécanique : c'est bien à cette version que je m'intéresse, et non aux raisons pour lesquelles différents agents ont cette même version.

Pour finir, notons que l'analyse de la section 4 a mis en évidence que le problème que l'hypothèse des versions est destinée à traiter est au cœur des réflexions de plusieurs savants et philosophes de la fin du dix-neuvième siècle, souvent considérés comme les ancêtres de la philosophie des sciences dans la tradition des empiristes logiques. En effet, soucieux de clarifier les fondements et la signification de la mécanique, Mach, ainsi que Hertz, reconnaissent explicitement que la compréhension d'une théorie est toujours le fait d'un agent sensible aux capacités cognitives limitées (pour le dire en termes modernes), et prennent acte de la difficulté conséquente d'identifier le contenu objectif d'une théorie. Sans considérer que leur autorité suffise à offrir un argument en faveur de la perspective que je défends, je tiens toutefois à souligner la parenté entre les problèmes qui les occupent et ceux qui m'intéressent ici, car ce sont précisément ces problèmes – ceux de la compréhension des agents individuels – que les empiristes logiques, et la tradition qui s'en est suivie, ont cherché à évacuer en tentant de définir le contenu objectif des connaissances scientifiques, indépendamment des esprits dont elles sont le fruit (ces approches, que j'appelle « formelles », sont examinées au chapitre 5 du présent chapitre).

Chapitre 4

Que signifie « comprendre la
mécanique classique » ?



$$\begin{aligned}\frac{dp}{dt} &= \Sigma_i \frac{dX_i}{dt} \frac{\partial L}{\partial q_i} + \Sigma_i X_i \frac{d}{dt} \left(\frac{\partial L}{\partial q_i} \right) \\ &= \Sigma_i Y_i \frac{\partial L}{\partial q_i} + \Sigma_i X_i \frac{\partial L}{\partial q_i} = 0 \ .\end{aligned}$$

Dessin de J.J. Sempé, *Le Petit Nicolas*.

Ce chapitre conclusif est destiné à tirer des analyses proposées aux chapitres 2 et 3 quelques considérations sur la notion de compréhension. Mon but est à la fois de dresser un bilan de ces analyses et de soulever d'autres questions qui leur sont liées, dont certaines seront abordées dans la suite de ce travail. Ce chapitre a donc un caractère programmatique : je me contente, ici, d'esquisser ces questions, et d'annoncer les chapitres qui les approfondiront.

Aborder l'étude de l'activité théorique du point de vue des agents implique, comme je l'ai mentionné à plusieurs reprises, de s'intéresser à la manière dont ils *comprennent* les hypothèses théoriques auxquelles ils sont confrontés, que ce soit au cours de l'apprentissage ou de la recherche. La première définition que j'ai donnée de la notion de version, avant son examen plus approfondi, consistait ainsi à décrire la version d'un agent comme la compréhension qu'il a d'une théorie. Comme je l'ai suggéré, la notion de version permet à la fois de distinguer différentes *manières* de comprendre une théorie, et d'éclairer les différents *degrés* auxquels on peut comprendre une théorie : la version d'un profane est à la fois moins robuste et moins développée que celle d'un expert. Dans ce chapitre, je me propose d'approfondir le sens de ce que j'entends ici par l'expression « comprendre une théorie ».

Dans un premier temps (section 1), je me propose d'examiner les compétences que l'on exige de la part d'un agent pour affirmer qu'il comprend (plus ou moins bien) une théorie ; on va voir qu'elles correspondent aux deux fonctions – représentative et inférentielle – des théories. Cela me permet, dans un deuxième temps (section 2), de soulever la question de l'intelligibilité d'une théorie, c'est-à-dire de ce qui fait qu'une théorie est susceptible d'être comprise par les agents, au sens défini dans la section qui précède. Autrement dit, après m'être intéressée à ce qu'un agent doit être capable de faire pour affirmer qu'il comprend une théorie, j'examine les caractéristiques qu'une théorie doit avoir pour qu'il soit possible aux agents de satisfaire ces exigences.

1 Compétences représentationnelles et computationnelles

Que veut-on dire, quand on affirme qu'un individu comprend une théorie ? Si l'on définit une théorie comme un ensemble d'hypothèses (ou de principes) permettant de prédire et d'expliquer les phénomènes empiriques, il semble que la bonne compréhension d'une théorie implique d'avoir une vue claire de l'architecture logique de ses différents principes et de la manière dont ils sont reliés au monde empirique. Sans remettre en cause une telle définition, je me propose, dans cette section, d'en préciser le sens. En réexaminant certains aspects du projet de Hertz (1894), et en proposant une rapide analyse critique de certains propos de Duhem (1914), j'opposerai deux conceptions de la compréhension d'une théorie. Selon la première, que j'examine dans un

premier temps (section 1.1), une compréhension parfaite impliquerait de s'abstraire de toute perspective particulière et d'en avoir vue globale et surplombante. La seconde conception repose sur le présupposé – qui motive l'ensemble de ce travail – selon lequel la compréhension d'une théorie est toujours *située*. Je montrerai, dans un deuxième temps (section 1.2), que la compréhension ainsi conçue implique la mise en œuvre de deux compétences, l'une permettant d'effectuer des calculs (compétence *computationnelle*) et l'autre permettant de se représenter le monde au moyen des concepts de la théorie (compétence *représentationnelle*).

1.1 La compréhension abstraite : Hertz et Duhem

J'ai signalé, au chapitre 3 (section 4.2.3), qu'il existe une tension, au sein même du projet de Hertz (1894), entre la défense d'une *Bildtheorie*, selon laquelle les individus raisonnent au moyen d'images mentales permettant de prédire les phénomènes, et la recherche d'une présentation des principes de la mécanique débarrassée de tout effet de perspective. Cette tension correspond au conflit de deux idéaux d'intelligibilité différents, sur lesquels je souhaite revenir ici.

La présentation parfaite des principes, visée comme un idéal par Hertz, consisterait à expliciter l'architecture logique de la mécanique de manière à offrir une vue *globale* et *surplombante* de tout l'édifice, des liens logiques entre ses principes, et de la manière dont l'ensemble est relié aux phénomènes. Cela permettrait de se dégager de la perspective particulière dans laquelle on se trouve pour en saisir le contenu, et de débarrasser ce dernier de tout appareil inutile. Au principe même de cet idéal se trouve l'idée selon laquelle on devrait, au moins en principe, pouvoir distinguer clairement entre ce qui relève du contenu d'une théorie et ce qui relève de sa formulation et des représentations qui lui sont associées dans l'esprit des individus. Comme je l'ai montré, cette visée d'une compréhension abstraite et sans perspective entre en conflit avec le constat, dressé par Hertz, que toute compréhension est inévitablement *située*, et s'appuie sur la manipulation de représentations particulières.

Ce dernier constat – celui de la nécessité, pour comprendre, de raisonner au moyen de représentations particulières – est précisément ce que rejette Pierre Duhem (1914, I, iv), dans ses propos sur l'opposition, inspirée de Pascal¹, entre les esprits abstraits (que Duhem attribue aux Français et aux Allemands) et les esprits amples (qu'il attribue aux Anglais). Comme je vais le montrer, en prétendant distinguer deux types d'esprits, Duhem exprime en fait une opposition entre deux définitions de la compréhension et des raisonnements qu'elle implique. Les exposer brièvement ici me

¹Pascal (1670) distingue entre l'« esprit de géométrie » et l'« esprit de finesse » (art. premier, pensées 1 et 2 dans l'édition Brunschvicg).

permettra de préciser la thèse que j'ai défendue tout au long de cette partie.

Rappelons d'abord la description faite par Duhem des esprits abstraits et des esprits amples. Les esprits abstraits sont ceux, selon Duhem, chez qui « la faculté de concevoir des idées abstraites et d'en raisonner est plus développée que la faculté d'imaginer des choses concrètes. » Ils « conçoivent sans effort une idée que l'abstraction a dépouillée de tout ce qui exciterait la mémoire sensible ; ils saisissent clairement et complètement le sens d'un jugement reliant de telles idées ; ils sont habiles à suivre, sans lassitude ni défaillance, jusqu'à ses dernières conséquences, un raisonnement qui prend pour principes de tels jugements. ». En revanche, ils sont incapables de « rendre présents aux yeux de l'imagination un très grand nombre d'objets, de telle façon qu'ils soient saisis tous à la fois » (Duhem, 1914, p. 78). Cette dernière aptitude caractérise les esprits amples, ou *imaginatifs*, qui sont capables d'imaginer et d'embrasser une multitude de détails concrets ; en revanche, ces esprits, selon Duhem, ne parviennent à s'exercer que sur des objets qui « tombent sous les sens, qui se touchent ou qui se voient. » Ces esprits ont besoin, « pour concevoir, du secours de la mémoire sensible ; l'idée abstraite, dépouillée de tout ce que cette mémoire peut figurer, leur semble s'évanouir comme un impalpable brouillard » (p. 79).

Si ces définitions opposent deux types d'esprits qui peuvent tous deux être « vigoureusement développés » (p. 79) et qui ont chacun une puissance qui leur est propre, Duhem juge explicitement que les esprits abstraits sont supérieurs, car ils sont seuls capables de saisir la « théorie physique abstraite », sans avoir recours, comme les physiciens anglais que Duhem critique violemment (pp. 99-109), à des « modèles mécaniques ». Je reviendrai plus loin sur la notion de modèle mécanique ; je voudrais insister, pour l'instant, sur la chose suivante : ce que Duhem oppose ici n'est pas, comme une lecture trop rapide pourrait le faire penser, d'une part la capacité à utiliser le langage mathématique d'une théorie, et d'autre part la préférence pour des « images » comme, par exemple, des schémas ou des diagrammes, ou même comme les images que suggèrent les analogies ou les métaphores auxquelles les scientifiques ont parfois recours (par exemple, l'analogie entre les molécules d'un gaz et les boules d'un jeu de billard en théorie cinétique des gaz).

Pour Duhem, les équations, tout comme les diagrammes, font partie de l'« habillage » concret de la théorie dont les esprits amples ont besoin, mais dont les esprits abstraits se passent sans difficulté. Reprenant la distinction classique entre l'entendement, siège du raisonnement, et l'imagination, il oppose les *raisonnements* logiques qui caractérisent la « méthode purement abstraite et déductive » que l'esprit abstrait peut conduire sur les notions abstraites, et le *calcul* propre à l'esprit ample, « où la faculté d'imaginer [a] plus de part que le pouvoir de raisonner ». Les mathématiques, en tant qu'elles ont recours à des symboles comme les nombres et à des règles de

manipulation de ces symboles, sont le fait, pour Duhem, des esprits amples :

Au lieu de traiter directement des notions abstraites qui les occupent, de les considérer en elles-mêmes, ils profitent de leurs propriétés les plus simples pour les représenter par des nombres, pour les *mesurer* ; alors, au lieu d'enchaîner dans une suite de syllogismes les propriétés de ces notions elles-mêmes, ils soumettent les nombres fournis par les mesures à des manipulations opérées suivant des règles fixes, les règles de l'Algèbre ; au lieu de déduire, ils *calculent*. Or, cette manœuvre des symboles algébriques qu'on peut, dans la plus large acception du mot, nommer le *calcul*, suppose, chez celui qui la crée comme chez celui qui l'emploie, bien moins la puissance d'abstraire et l'habileté à conduire par ordre ses pensées, que l'aptitude à se représenter les combinaisons diverses et compliquées qui se peuvent former avec certains signes visibles et dessinables, à voir d'emblée les transformations qui permettent de passer d'une combinaison à l'autre ; l'auteur de certaines découvertes algébriques, un Jacobi par exemple, [...] ressemble [...] au joueur qui conduit à une victoire assurée la tour ou le cavalier. (Duhem, 1914, pp. 89-90)

La faculté de manipuler des symboles ne s'oppose pas, pour Duhem, à l'imagination et à ce qu'on appelle parfois la « visualisation » : elle en fait partie. Les déductions purement logiques que les esprits abstraits sont capables de faire ne s'apparentent pas à la manipulation de symboles algébriques.² Ces symboles, et leurs règles de manipulation, pallient ainsi, selon Duhem, le défaut de capacité d'abstraction des esprits amples. Les esprits abstraits peuvent s'en passer ; leurs raisonnements ne s'appuient pas sur des *représentations*, et ne consistent pas en des *calculs* ; ils saisissent directement la structure logique de la théorie, derrière toute formulation particulière.³ La « théorie physique abstraite », en tant qu'objet des esprits abstraits, n'est donc ni outil de représentation, ni outil d'inférence.

Je ne discuterai pas la question de savoir si, comme le prétend Duhem, il existe des esprits (humains) capables de conduire ce qu'il appelle ici des « raisonnements »,

²En termes modernes, une telle manipulation de symboles algébriques n'est pas, selon Duhem, *purement syntaxique* : en tant qu'elle implique, comme il le souligne, la faculté d'imagination, les symboles que l'on manipule sont déjà, d'une certaine manière, interprétés. On retrouvera ces considérations lors de l'examen de la conception empiriste logique des théories, au chapitre 5, section 1.5.2 (en particulier page 216).

³« Pour un Français ou pour un Allemand, une théorie physique est essentiellement un système logique, des déductions parfaitement rigoureuses unissent les hypothèses sur lesquelles repose la théorie aux conséquences qu'on en peut tirer et qu'on se propose de comparer aux lois expérimentales ; si le calcul algébrique intervient, c'est seulement pour rendre moins lourde et plus maniable la chaîne de syllogismes qui doit relier les conséquences aux hypothèses ; mais dans une théorie saine-ment constituée, ce rôle purement auxiliaire de l'Algèbre ne doit jamais se laisser oublier ; il faut qu'on sente, à chaque instant, la possibilité de remplacer le calcul par le raisonnement purement logique dont il est l'expression abrégée [...]. » (Duhem, 1914, pp. 112-113)

qui ne s'appuient aucunement sur des représentations.⁴ Je m'appuierai plutôt sur les propos de Duhem pour montrer, *a contrario*, qu'une théorie scientifique, *en tant qu'outil d'inférence* – et donc clairement distincte de la « théorie physique abstraite » de Duhem – ne remplit ce rôle qu'en tant qu'elle fournit des représentations sous une forme particulière aux agents. Ce que les propos de Duhem permettent de mettre en évidence, c'est que c'est l'« habillage concret » de la théorie, c'est-à-dire sa formulation particulière, qui en fait un outil d'inférence, et que ces inférences consistent en la manipulation de représentations particulières.

1.2 Les deux dimensions de la compréhension située

Le premier élément de définition de la compréhension d'une théorie, donné au début de cette section, mérite donc d'être précisé : si comprendre une théorie, c'est avoir une vue claire de l'articulation de ses principes et de la manière dont ils sont reliés aux phénomènes empiriques, cette compréhension est toujours située, et repose sur la manipulation de représentations particulières. En tant que telle, elle implique une compétence *inférentielle* ou *computationnelle* (c'est-à-dire permettant d'effectuer des *calculs*) : la connaissance de cette articulation logique s'apparente plus à un savoir-faire et une maîtrise pratique qu'à une connaissance abstraite et sans perspective. Elle consiste en effet à *être capable* de déduire les différents principes les uns des autres, et à savoir leur donner la forme adéquate qui permettra de les manipuler de façon à en tirer les solutions d'un problème particulier. Cette compétence peut être plus ou moins développée, et s'acquiert manifestement par l'exercice : selon que l'on maîtrise plus ou moins bien les règles du calcul différentiel, par exemple, on est plus ou moins capable de manipuler les équations newtoniennes du mouvement, et de déduire les différents principes de la mécanique les uns des autres.

Cette compétence *computationnelle* est indissociable, pour la compréhension d'une théorie *empirique*, d'une compétence que j'appelle *représentationnelle*. En effet, la manipulation des représentations fournies par la théorie – par exemple, des équations différentielles – implique, certes, une capacité à conduire des calculs selon les lois de la déduction mathématique, mais aussi à en comprendre la signification empirique. Comme on l'a vu au chapitre 2 (section 1.3.1), la connaissance de la deuxième loi de

⁴L'idée d'une théorie scientifique conçue comme un système logique a été développée par les empiristes logiques, qui lui ont donné un sens précis, que j'analyserai au chapitre 5. La question de savoir si, comme le suggère Duhem, il est possible d'*utiliser* ce système pour opérer des déductions, sans que cela requière la manipulation de symboles, peut être rapprochée de la question de la possibilité du formalisme pur. La plupart des philosophes des mathématiques, aujourd'hui, la rejettent, ce qui semble aller dans le sens de mon propos (merci à Paul Humphreys pour m'avoir suggéré ce lien entre mon analyse et la critique du formalisme pur).

Newton – et du calcul différentiel – ne suffit pas à résoudre des problèmes. Il faut aussi, pour prédire le comportement d'un pendule, savoir identifier les forces qui s'y appliquent de façon à écrire l'équation sous une forme adéquate – qui, dans le cas du pendule, fait intervenir le sinus de l'angle d'oscillation, et est assez différente de celle sous laquelle la deuxième loi de Newton est habituellement exprimée. Il faut ensuite savoir identifier les conditions initiales du système afin de résoudre cette équation. Cela implique de faire correspondre les symboles encore abstraits de la loi à une situation concrète.⁵ Autrement dit, cela implique d'être en mesure de *se représenter les phénomènes* au moyen des concepts fournis par la théorie – en l'occurrence, celui de force – afin de lui faire dire quelque chose à propos du monde empirique. Dans les termes de Thomas Kuhn, je comprends la mécanique newtonienne si je suis capable de *voir des situations concrètes comme* des situations newtoniennes (voir par exemple Kuhn, 1970b, pp. 169-170). Cette compétence représentationnelle, tout comme la compétence computationnelle, est susceptible de degrés, et s'acquiert par l'exercice : comme le souligne Kuhn (Kuhn, 1969, p. 470), les exercices que l'on trouve à la fin des manuels de mécanique sont destinés à développer la capacité des étudiants à reconnaître des situations de plus en plus diverses comme des situations newtoniennes. Ces exercices servent, précisément, à transformer la compréhension abstraite que les étudiants ont acquise en apprenant leur leçon en une compréhension située et concrète, leur permettant d'utiliser la théorie pour résoudre des problèmes.⁶

Ces deux dimensions de la compréhension d'une théorie – représentationnelle et computationnelle – sont indissociables. En effet, c'est en tant que je comprends ce qu'elle dit des phénomènes, c'est-à-dire en tant qu'elle en est une représentation, que je suis en mesure d'écrire une équation *sous la forme adéquate*. C'est le sens que je propose de donner à l'affirmation de Feynman, selon laquelle la compréhension d'une équation, pour un physicien, relève d'un autre type de compétence que la seule compréhension mathématique :

Ce que cela signifie réellement de comprendre une équation – c'est-à-dire, pas seulement dans un sens strictement mathématique – a été décrit par Dirac. Il a dit : « Je

⁵L'analyse des propos de Duhem a permis de mettre en évidence le fait que la compétence computationnelle, en elle-même, implique déjà une compétence représentationnelle : comme je l'ai souligné, la manipulation de symboles algébriques semble reposer en elle-même sur une faculté d'imagination ; si c'est le cas, les symboles algébriques, comme je l'ai suggéré, ne sont pas à strictement parler abstraits. Ce que je souligne ici est que, dans le cas des sciences empiriques, cette manipulation est indissociable de leur interprétation *empirique*, c'est-à-dire de l'attribution de référents empiriques aux symboles que l'on manipule. C'est donc pour ainsi dire d'une « deuxième couche » interprétative qu'il s'agit ici.

⁶Le chapitre 6 de ce travail est en partie consacré à l'analyse des propos de Kuhn sur l'apprentissage scientifique.

comprends ce qu'une équation signifie si j'ai un moyen de me figurer les caractéristiques de sa solution sans avoir besoin de la résoudre effectivement. » Ainsi, si l'on a un moyen de savoir ce qui devrait arriver dans des circonstances données sans résoudre effectivement les équations, alors on « comprend » les équations, en tant qu'elles sont appliquées à ces circonstances. Une compréhension physique est une chose totalement non-mathématique, imprécise, et inexacte, mais absolument nécessaire pour un physicien. (Feynman *et al.*, 1963, II, 2-1)

La dimension computationnelle de la compréhension d'une équation n'est pas purement mathématique – elle n'est pas seulement la connaissance du calcul différentiel qui permet la manipulation de symboles abstraits⁷ – au sens où elle implique l'utilisation de cette équation pour représenter une situation physique. Elle implique donc de comprendre *ce que dit cette équation*, et cette compréhension guide, comme le suggère Feynman, les calculs que le physicien effectue.⁸

Pour terminer, notons que, non seulement cette compréhension *physique* guide les calculs effectués en vue de résoudre des problèmes, mais on peut également supposer qu'elle guide l'invention théorique elle-même, en tant qu'elle consiste à déduire de nouveaux principes et à développer l'appareil mathématique d'une théorie. Feynman ajoute d'ailleurs, à propos de la croyance selon laquelle la compréhension mathématique parfaite [*inside out*] d'une équation suffit à donner une compréhension parfaite de la physique, que

Cela ne marche pas comme ça. Les mathématiciens qui étudient la physique dans cette perspective – et il y en a eu beaucoup – apportent généralement bien peu de contributions à la physique, et, en fait, aux mathématiques. Leur échec est dû au fait que les situations physiques concrètes dans le monde réel sont si compliquées qu'il est nécessaire d'avoir une compréhension bien plus large des équations. (Feynman *et al.*, 1963, II, 2-1)

Ce que semble vouloir dire Feynman, c'est que le développement d'une théorie, même quand il consiste, comme dans le cas de la mécanique classique, en l'exploration des conséquences déductives d'un ensemble de principes, ne se réduit pas à l'application pure et simple des règles de déductions mathématiques. Il est guidé par la volonté de tirer des conséquences *physiques* de ce principe. C'est ce qui permet de dire, comme je l'ai proposé au chapitre 3 (section 5.3, page 157), qu'un changement formel (mathématique) peut modifier la signification *physique* d'une théorie.

Cela invite finalement à reconsidérer l'affirmation selon laquelle Lagrange se comporte en mathématicien (voir chapitre 3, page 114) : si, de son propre aveu, il ne fait

⁷Voir ci-dessus, note 5.

⁸L'analyse de ce que j'appelle ici la dimension « computationnelle » de la compréhension, et de la relation qu'elle entretient avec ce que j'appelle sa dimension « représentationnelle », sera un des objets principaux du chapitre 6 (en particulier section 3.1.1).

qu'apporter une modification formelle à la mécanique, on peut légitimement affirmer que son enquête mathématique est guidée par une authentique *compréhension physique* (ou, si l'on veut, *intuition physique*). En ce sens, son travail requiert bien plus que des talents de mathématicien – si du moins l'on réduit, comme le fait Feynman, les talents du mathématicien à des compétences « étroitement » computationnelles, c'est-à-dire n'impliquant que la manipulation de symboles abstraits.⁹

Dans cette section, j'ai opposé deux conceptions différentes de la compréhension : celle, présentée comme un idéal par Hertz et vigoureusement défendue par Duhem, selon laquelle la compréhension d'une théorie implique de s'abstraire de toute perspective particulière et d'en saisir le contenu sans la médiation de représentations mentales, et celle, qui sous-tend l'ensemble du présent travail, selon laquelle la compréhension est toujours située et repose sur la manipulation de représentations particulières. Adopter cette dernière perspective invite à se demander, comme je propose de le faire dans la section suivante, quelles exigences une théorie doit remplir pour être intelligible, c'est-à-dire pour permettre aux individus de manipuler ces représentations.

2 Qu'est-ce qui fait qu'une théorie est intelligible ?

Dans la mesure où la compréhension d'une théorie implique la capacité, de la part des agents, à manipuler des représentations afin de tirer des inférences à propos des phénomènes, l'intelligibilité d'une théorie pour un agent doit dépendre de deux ensembles de facteurs : d'une part, de certaines caractéristiques de la théorie qui permettent de la présenter sous une forme utilisable par l'agent, et d'autre part des compétences et des capacités cognitives de l'agent. Autrement dit, l'intelligibilité d'une théorie semble dépendre de la relation entre la forme sous laquelle elle est présentée et des capacités cognitives des agents.¹⁰

Ces considérations sont au cœur de certaines réflexions exprimées par Boltzmann (1905) dans le cadre de sa *Bildtheorie*. Pour Boltzmann comme pour Hertz, les images que nous fournissent les théories doivent être en accord avec les lois de la pensée. Cependant, comme je l'ai signalé au chapitre 3 (note 56), il existe un désaccord important entre Hertz et Boltzmann, concernant ce qu'il faut entendre par « lois de

⁹Ici encore, l'adjectif « abstrait » n'est pas à comprendre au sens duhémien, mais au sens de « non encore interprété *empiriquement* ».

¹⁰La relation entre la forme particulière des représentations théoriques et les capacités cognitives de leurs utilisateurs, dont j'esquisse l'analyse dans cette section, fera l'objet d'un examen approfondi au chapitre 7, sections 2 et 3.

la pensée ». Dans cette section, je présente brièvement ce désaccord¹¹ en montrant les conséquences de la conception de Boltzmann pour une réflexion sur l'intelligibilité des théories (section 2.1). Je montre ensuite (section 2.2) comment la définition de l'intelligibilité qui se dégage de cet examen permet d'éclairer certaines questions sur la compréhension que différents individus, de l'expert au profane, peuvent avoir d'une théorie – sur leur *version* de cette théorie.

2.1 Boltzmann et l'évolution des lois de la pensée

Pour Hertz, affirmer que les images doivent être en accord avec les lois de la pensée signifie qu'elles doivent être conformes aux lois logiques, qu'il considère universelles et éternelles ; en conséquence, la question de ce qu'il appelle l'« admissibilité logique » (voir chapitre 3, page 136) d'une image peut être tranchée objectivement et pour toujours. C'est la raison pour laquelle une présentation parfaite des principes de la mécanique doit s'efforcer, autant que possible, de les débarrasser de toutes les « relations superflues » dues à notre esprit et à son « mode de reproduction des objets ». Le fait qu'une image porte la marque de l'esprit qui la produit est, pour Hertz, un signe d'obscurité, et non de clarté.

Pour Boltzmann, en revanche, le fait qu'il soit indispensable d'être en mesure de se former une image intuitive de nos théories n'est pas seulement dû à une limitation de notre esprit, c'est la condition même pour qu'une théorie soit intelligible – pour qu'elle joue son rôle d'outil de représentation et d'inférence et, ce faisant, pour qu'elle nous permette de comprendre les phénomènes. Les « lois de la pensée » auxquelles une théorie doit se conformer ne sont donc pas des lois logiques *a priori* et universelles, mais bien la manière particulière dont nous nous représentons les phénomènes. Or, selon Boltzmann, les lois de la pensée, ainsi conçues, ne sont pas éternelles : à l'idée kantienne du caractère *a priori* des lois logiques et des formes de l'intuition, Boltzmann oppose une conception que l'on a souvent qualifiée de darwinienne, selon laquelle les lois de la pensée évoluent, au fur et à mesure des progrès de la physique et de la connaissance en général. Autrement dit, si les images que nous fournissent les théories portent la marque de notre esprit, c'est aussi parce que notre capacité à nous former de telles images est un produit de l'adaptation. En conséquence, selon Boltzmann, ce qui est intelligible aujourd'hui – ce que les agents cognitifs de l'époque de Boltzmann sont capables d'utiliser comme outil de représentation et d'inférence – pourrait cesser de l'être demain, si la découverte de nouveaux phénomènes empi-

¹¹Pour plus de détail sur ce désaccord, et sur la définition de la compréhension qui se dégage des propos de Boltzmann, voir (de Regt, 1999). Pour la philosophie de Boltzmann, voir les références indiquées chapitre 3 (note 54, page 134).

riques et les progrès de la science nous conduisaient à inventer de nouvelles formes de représentations.

Notons que ces réflexions de Boltzmann trouvent un écho, quelques décennies plus tard, dans les propos de certains fondateurs de la mécanique quantique. Cette dernière, en remettant en cause la validité universelle des concepts classiques d'espace et de temps, bouleverse en effet les critères habituels d'intelligibilité des théories fournis par la physique classique. Son caractère « non intuitif », qui interdit de se représenter les objets physiques dans l'espace et le temps, conduit ainsi Heisenberg, entre autres, à redéfinir le « caractère intuitif » d'une théorie, « en le détachant de sa référence à la région de notre expérience habituelle » (Chevalley, 1998, p. 159, note 3). Comme Boltzmann, il propose, contre l'idée kantienne de leur validité *a priori*, une conception évolutionniste des formes de l'intuition :

Ces formes de l'intuition [...] ont fait leurs preuves *dans le commerce de l'homme avec le monde*¹², et elles doivent leur validité précisément à ce fait.

[...] le fait que nous ne pouvons faire absolument aucune expérience autrement que dans ces formes de l'intuition [...] ne légitime pas la supposition que les formes de l'intuition restent inchangées à toutes les époques. L'existence de la théorie de la relativité montre au contraire déjà que nous pouvons et que nous devons ici aussi apprendre autrement, et l'on peut penser qu'à des époques futures la manière de faire des expériences et de les agencer sera dès le début différente de la nôtre. Le biologiste Lorenz a soutenu qu'il faut considérer les formes de l'intuition comme des « schèmes innés » et les comparer à l'instinct d'action des animaux. Cette conception explique, d'une part, pourquoi ces formes de l'intuition forment pour nous la présupposition nécessaire de toute expérience [...]. D'autre part, le fait que ces formes de l'intuition « fassent leurs preuves » dans la réalité devient une condition de notre existence : seuls les êtres dont les formes de l'intuition s'ajustent bien à leur monde ambiant peuvent subsister dans le combat pour l'existence. (Heisenberg, 1942, pp. 356-357)¹³

Autrement dit, que la mécanique quantique semble inintelligible à une certaine époque ne signifie pas qu'elle soit destinée à le rester. Le standard d'intelligibilité fourni par les théories classiques ne doit pas, en ce sens, être considéré comme universel et éternel.¹⁴

Il n'en reste pas moins que, à l'époque de Boltzmann, les modèles classiques sont les seuls qui soient intelligibles. Sur la base de sa conception évolutionniste des lois

¹²Je souligne.

¹³Je cite la traduction de Catherine Chevalley, indiquée en bibliographie.

¹⁴Sur la manière dont les fondateurs de la théorie quantique cherchent à rendre compte de l'impossibilité de fournir un modèle classique de cette théorie, voir l'introduction et l'entrée « Images / Représentations intuitives » du glossaire de l'édition de Catherine Chevalley (1991) aux écrits de Niels Bohr (1961), ainsi que son introduction (Chevalley, 1998) à (Heisenberg, 1942).

de la pensée, Boltzmann défend en effet, avec des physiciens anglais comme Maxwell (1890) et Kelvin (Thomson, 1884)¹⁵, que Duhem (1914, pp. 99-109) attaque vivement, l'idée selon laquelle une théorie n'est intelligible – et donc ne fournit une image explicative des phénomènes – que si l'on peut en présenter un *modèle mécanique*. Cela signifie qu'une théorie doit nous permettre de nous représenter les phénomènes *comme* des phénomènes mécaniques ; par exemple, elle doit pouvoir être exprimée, et utilisée pour représenter les phénomènes, au moyen d'équations qui ont la *forme* des équations de la mécanique.¹⁶

Je reviendrai, au chapitre 5 (sections 1.5 et 1.6), sur la notion de modèle mécanique, et sur son lien avec la question de l'intelligibilité des théories. Il convient cependant, dès à présent, de souligner un point important. Comme je l'ai dit, Boltzmann admet qu'une théorie qui n'est pas intelligible à son époque – parce que l'on ne peut pas en fournir un modèle mécanique – est susceptible de le devenir un jour. En conséquence, sa préférence pour les modèles mécaniques ne s'accompagne pas chez lui de la défense de l'hypothèse atomiste, selon laquelle le monde *est* effectivement constitué de matière et de mouvement. Affirmer que l'on ne comprend une théorie – et les phénomènes qu'elle explique – que si elle est conforme à notre manière de nous

¹⁵ Ainsi, Thomson écrit : « Il me semble que le vrai sens de cette question : Comprendons-nous ou ne comprenons-nous pas tel sujet de Physique ? est celui-ci : Pouvons-nous construire un modèle mécanique correspondant ? » (Thomson, 1884, p. 131, cité par Duhem, 1914, p. 102), et, plus loin : « Je ne suis jamais satisfait tant que je n'ai pu construire un modèle mécanique de l'objet que j'étudie ; si je puis faire un modèle mécanique, je comprends ; tant que je ne puis pas faire un modèle mécanique, je ne comprends pas ; et c'est pourquoi je ne comprends pas *la* théorie électromagnétique de la lumière. Je crois fermement en *une* théorie électromagnétique de la lumière [...] » (Thomson, 1884, p. 270, cité par Duhem, 1914, p. 103).

¹⁶ Notons que l'on peut s'interroger sur l'univocité de la notion de modèle mécanique, dans la mesure où, comme on l'a vu, les équations de la mécanique peuvent elles-mêmes être présentées sous plusieurs formes différentes. Maxwell, à propos de la méthode de Lagrange qui consiste à éliminer les forces de contrainte exercées sur un système, écrit ainsi : « En suivant les étapes de cette élimination, l'esprit doit exercer sa capacité de calcul, et doit donc se prémunir contre l'intrusion d'idées dynamiques. Notre but, au contraire, est de cultiver nos idées dynamiques. Nous profitons donc des travaux des mathématiciens, et retraduisons leurs résultats depuis le langage de l'analyse (calculus) vers le langage de la dynamique, de telle sorte que nos mots puissent appeler une image mentale, non pas de quelque processus algébrique, mais d'une propriété des corps en mouvement. » (Maxwell, 1920, pp. 123-124) Si Lagrange a opéré un progrès formel remarquable, ses équations ont perdu en intelligibilité pour un esprit mécaniste comme celui de Maxwell. Autrement dit, il semble que les équations de Lagrange n'offrent pas à strictement parler, selon Maxwell, un modèle mécanique des phénomènes. Cela s'accorde avec l'idée, que je défends, selon laquelle les différentes formulations de la mécanique peuvent être décrites comme différentes manières de la comprendre (comme différentes versions). Voir à ce sujet les réflexions de Jeremy Butterfield (2004), qui affirme que l'image de la « matière en mouvement », généralement attribuée à la mécanique classique, n'en est qu'un modèle possible. On retrouvera ce problème au chapitre 5 (section 1.6.5).

représenter les phénomènes, et si, par conséquent, on peut l'utiliser comme outil de représentation et d'inférence, n'implique pas que l'on croie que le monde est effectivement tel que cette théorie le représente.¹⁷ Autrement dit, les propos de Boltzman peuvent être interprétés comme l'expression d'une thèse portant sur la nature des théories et de l'activité théorique plutôt que sur le monde.

2.2 Des standards d'intelligibilité aux versions individuelles

On vient de voir que, pour Boltzmann, en vertu même de ce que j'appelle la double fonction représentative et inférentielle des théories, leur intelligibilité dépend à la fois de la forme sous laquelle elles peuvent être exprimées et de la nature même du raisonnement humain. Ses considérations portent sur l'intelligibilité des théories, à une époque donnée, pour *tous les scientifiques* ; autrement dit, les modèles mécaniques sont des *standards d'intelligibilité* à l'échelle large de la communauté scientifique de la fin du dix-neuvième et du début du vingtième siècles. Je n'approfondirai pas, dans ce travail, l'analyse du problème de la variabilité historique des standards d'intelligibilité¹⁸ ; en revanche, je souhaite tirer de ces considérations des conséquences pour une analyse de la compréhension à l'échelle beaucoup plus restreinte de l'individu.

J'ai proposé, au chapitre 3 (section 2.1.1), de décrire l'apprentissage d'une théorie par un agent – l'approfondissement de sa compréhension de cette théorie – comme un développement et un renforcement de la version qu'il en a. J'ai suggéré que cela consistait à la fois à développer ses capacités de calcul, permettant de rendre les liens inférentiels entre différents principes et équations plus solides, et à apprendre à se représenter des phénomènes de plus en plus divers et nombreux au moyen des concepts de la théorie. En ce sens, ce processus consiste à se familiariser progressivement avec les différentes formes sous lesquelles une théorie peut être présentée. Alors que les équations lagrangiennes sont inintelligibles au profane, l'expert les manie avec une aisance qui lui permet de se représenter les phénomènes par leur moyen. Comprendre une théorie reviendrait ainsi à faire évoluer, par l'exercice, ses propres habitudes de raisonnement de manière à être capable d'utiliser les représentations fournies par cette théorie.¹⁹ De même que les lois de la pensée de toute la communauté scientifique d'une

¹⁷ Je reviendrai sur ce problème au chapitre 5, section 1.5.3, et au chapitre 6, pages 292-295.

¹⁸ Sur l'aspect contextuel des standards d'intelligibilité (dans une perspective différente de celle de Boltzmann), et ses conséquences pour la notion de compréhension, voir (Friedman, 1974; de Regt et Dieks, 2005). Voir aussi les réflexions de Hanson (1963), qui suggère qu'une théorie devient intelligible à mesure qu'elle prouve son pouvoir prédictif.

¹⁹ À propos de l'idée selon laquelle la compréhension des représentations théoriques – en particulier celles qui sont exprimées en langage mathématique – est essentiellement le fruit de l'habitude, citons la réponse que Johann von Neumann aurait faite à un jeune physicien qui lui aurait dit ne pas

époque évoluent, selon Boltzmann, au fur et à mesure que la physique progresse, on peut décrire le processus par lequel ce qui est incompréhensible au débutant le devient peu à peu comme un développement de ses capacités cognitives et de ses compétences computationnelles et représentationnelles par la manipulation de certaines formes d'équations.

En outre, j'ai souligné que la notion de version ne permet pas seulement de rendre compte de la différence de compréhension entre experts et profanes, mais également entre différents experts (voir chapitre 3, section 2.1.2). Pour reprendre un exemple que j'ai donné à ce moment-là, on peut supposer qu'un spécialiste de mécanique quantique, en raison de la familiarité qu'il a acquise avec les équations hamiltoniennes, se représente spontanément les phénomènes au moyen de ces équations, alors qu'un autre expert, plus habitué à manipuler les équations lagrangiennes, comprend plus aisément la mécanique sous ce formalisme. De plus, comme je l'ai également suggéré, selon que l'on préfère se représenter les phénomènes en termes de force ou en termes d'énergie, on n'attribuera pas exactement la même signification aux différents principes de la mécanique. L'analyse des versions exprimées au cours de l'histoire par différents savants et philosophes (chapitre 3, section 4) a renforcé cette hypothèse : selon leurs habitudes de raisonnement, leur domaine de spécialité, ainsi que leurs préférences et engagements théoriques, on peut affirmer que différents experts n'ont pas exactement la même version de la mécanique. Au delà des standards d'intelligibilité valables pour tous les scientifiques d'une communauté donnée, on peut ainsi supposer que les différents individus ne suivent pas exactement les mêmes processus inférentiels et ne manipulent pas exactement les mêmes représentations mentales ; en ce sens, sans que cela implique un jugement quant au *degré* de compréhension, on peut dire qu'on ne comprend pas la mécanique *de la même manière* selon la version que l'on en a – on ne met pas en œuvre, pour ainsi dire, les mêmes compétences représentationnelles et computationnelles.

Conclusion

Je n'ai fait, dans ce chapitre, que soulever un certain nombre de questions qui feront l'objet des analyses de la suite de ce travail. Ce dernier, dans son ensemble, vise en effet à tirer les conséquences du présupposé selon lequel les théories, en tant qu'outils de représentation et d'inférence, sont toujours comprises et utilisées depuis une perspective particulière. Le chapitre 5 est consacré à l'examen critique de cer-

comprendre la théorie de Hamilton-Jacobi : « Jeune homme, en mathématiques, on ne comprend pas ; on s'habitue, c'est tout. » (Cité dans Zukav, 1979, note p. 208)

taines conceptions des théories scientifiques nées de l'adoption du présupposé inverse, selon lequel le contenu d'une théorie doit pouvoir être identifié en faisant abstraction de son utilisation par les agents cognitifs. Cette analyse me permettra de fonder ma démarche sur le constat des limites internes de toute entreprise d'identification du contenu des théories qui ne prend pas en compte la perspective située de leurs utilisateurs.

Conclusion de la première partie

Récapitulons les problèmes abordés tout au long de cette première partie. Le but principal en était de remettre en question, par l'analyse de l'exemple canonique de la mécanique classique, la pertinence de la notion de théorie, telle qu'elle est habituellement conçue, pour une étude de l'activité scientifique des agents ; le second but en était d'introduire la notion de version et d'en montrer la fécondité.

Au chapitre 2, l'examen des différentes formulations de la mécanique, telles qu'elles sont enseignées et utilisées aujourd'hui, m'a permis de montrer que la conception traditionnelle des théories, qui en identifie le contenu à la clôture déductive d'un ensemble d'hypothèses, ne permet pas de rendre compte de leur utilisation comme outil d'inférence et de représentation. Quand on prête attention à cette double fonction, l'idée d'une équivalence entre les différentes formulations de la mécanique est remise en cause, et les contours de la notion de théorie se brouillent.

Prenant acte de cela, j'ai défini, au chapitre 3, la notion de version, destinée à aborder ce que l'on appelle généralement une « théorie » sous l'angle de la compréhension que les agents en ont. À la lumière de cette notion, j'ai proposé une analyse de plusieurs problèmes posés par les différentes présentations possibles (synchroniquement et diachroniquement) de la mécanique. J'ai, en particulier, montré que la distinction entre une différence (ou une avancée) formelle et une différence (ou une avancée) conceptuelle n'est pas aussi tranchée que les définitions logico-empiriques du contenu des théories le laissent entendre.

Le chapitre 4 a dressé un bilan de ces différents problèmes à la lumière d'une analyse de la notion de compréhension. Demander en quoi consiste « comprendre ce que dit la mécanique » pour un agent particulier implique manifestement de renoncer à l'idée selon laquelle « ce que dit la mécanique » peut être identifié sans la prise en compte du détail de l'utilisation des représentations pour tirer des inférences.

Au terme de cette analyse, deux questions se posent, auxquelles la deuxième partie de ce travail va répondre. La première découle d'une objection. J'ai montré, dans ce qui précède, l'échec d'un certain nombre de tentatives d'identification du contenu de la mécanique à un ensemble d'hypothèses. J'ai principalement montré

que la prise en compte de la compréhension individuelle, dans sa double dimension représentationnelle et computationnelle, implique de renoncer à la définition d'un tel contenu. On peut m'objecter que je n'ai prouvé ni l'impossibilité, en principe, d'une telle définition, ni, corrélativement, la pertinence de la prise en compte de la compréhension pour cette définition. La perspective dans laquelle j'ai abordé l'étude de la mécanique peut être décrite comme reposant sur une pétition de principe : étudier les différentes formes sous lesquelles elle a été exprimée conduit à mettre cette diversité en valeur et à douter de l'existence d'un contenu indépendant de la forme sous laquelle il est présenté.

Définir d'emblée les théories comme des outils de représentation et d'inférence – ce qui constitue le présupposé fondamental de l'ensemble de cette thèse – est une option que l'on peut contester. Il semble bien qu'il reste un sens de la notion de théorie qui autorise à affirmer que la mécanique est une seule et même théorie : ses différentes formulations ont les mêmes conditions de vérité empiriques, et jamais l'expérience ne nous contraindra à en choisir une. On est encore en droit de supposer que l'on peut, en principe, identifier le contenu de cette théorie (par delà ses formulations), et que l'échec jusqu'ici (en l'occurrence, jusqu'au début du vingtième siècle) ne signifie pas que cela est impossible.

Dans le chapitre 5, j'étudie ainsi la manière dont certains philosophes des sciences du vingtième siècle, héritiers des physiciens philosophes, ont élaboré et utilisé des outils logiques et mathématiques pour reconstruire formellement le contenu des théories scientifiques ; ces outils, selon eux, doivent permettre de montrer l'identité d'une théorie par delà ses différentes formulations. Au terme de ce chapitre, je conclurai que cette entreprise n'est pas seulement impossible ou vaine en pratique, mais qu'elle l'est (dans une certaine mesure) en principe.

Prenant acte de cela, je chercherai, dans les chapitres 6 et 7, à répondre à la deuxième question que l'on peut poser au terme du présent chapitre : de quels outils dispose-t-on pour analyser le détail de la représentation et de l'inférence scientifiques, dont j'ai montré qu'elles doivent faire l'objet d'une attention plus poussée ? Après avoir proposé quelques éléments de réponse à cette question, je les mettrai en œuvre dans une seconde étude de cas, qui occupera la troisième partie de ce travail.

Deuxième partie

Unités et outils d'analyse de l'activité théorique

Vue d'ensemble de la deuxième partie

L'analyse de l'exemple de la mécanique, dans la première partie de ce travail, a remis en question la pertinence de la notion classique de théorie pour une étude de l'activité théorique du point de vue des agents. Opérant un changement de perspective, j'ai introduit la notion de version, afin de rendre compte de l'importance de la manière dont les individus comprennent et utilisent différentes hypothèses théoriques dans leurs raisonnements.

La notion de version, telle que je l'ai définie, a un statut *hypothétique* : au chapitre 3, j'ai étudié différentes présentations de la mécanique *en supposant* qu'elles sont les expressions d'autant de versions de la mécanique. J'ai également suggéré que la version d'un agent, définie comme un ensemble de représentations mentales, se manifeste dans l'utilisation qu'il fait de représentations externes pour résoudre des problèmes ; ces représentations externes sont alors considérées comme la source et l'expression des représentations mentales qui constituent sa version.

En raison même de ce caractère hypothétique, et parce que nous n'avons pas d'accès direct aux représentations mentales des agents, la notion de version, en tant que telle, ne peut pas remplacer la notion de théorie comme *unité d'analyse* de l'activité scientifique. Par « unité d'analyse », j'entends désigner des éléments de la pratique scientifique que l'on peut isoler et dont la description et l'analyse permettent de rendre compte de la connaissance produite par l'entreprise scientifique.

Se pose alors une double question : d'une part, celle de la définition des *unités d'analyse* qui, à la place de celle de théorie, permettent d'étudier les détails de la représentation et de l'inférence scientifiques ; d'autre part, celle des *outils d'analyse* qu'il convient de leur appliquer. Répondre à cette double question est l'objet de cette deuxième partie. D'une part, je reviens sur les analyses disponibles de l'activité scientifique, dont j'approfondis la critique par une confrontation directe aux thèses de différents auteurs. D'autre part, je cherche à définir, par contraste avec ces approches, les unités et les outils permettant d'analyser l'activité théorique dans la perspective

que je défends.

Le chapitre 5 revient ainsi sur la notion classique de théorie, par une critique de ce que j'appelle les approches « formelles » de la connaissance scientifique. La première de ces deux approches est celle des empiristes logiques, héritiers des travaux du néo-positivisme du Cercle de Vienne. La seconde, appelée « conception sémantique », se présente comme une critique de l'approche des empiristes logiques. Elles reposent cependant toutes deux sur une conviction commune : selon elles, la diversité des formulations d'une théorie n'a pas de conséquence sur son contenu. Cette conviction fonde la démarche formaliste qui consiste à chercher dans les outils de la logique formelle et des mathématiques le moyen de reconstruire le contenu des théories par delà les différentes formulations sous lesquelles elles sont habituellement utilisées par les scientifiques. La différence majeure entre ces deux approches réside dans le choix de ces outils formels et dans le but qu'elles assignent à la reconstruction formelle des théories. Je chercherai à montrer que le présupposé commun de ces deux conceptions – selon lequel on peut, en principe, distinguer entre la forme et le contenu des théories – est responsable de plusieurs tensions *internes* au sein de ces deux approches. Ma critique vise, d'une part, à préciser ma démarche par contraste avec une perspective qui repose sur le présupposé inverse du mien ; d'autre part, elle vise à fonder mon approche sur le constat des limites internes de toute entreprise visant à étudier le contenu des connaissances scientifiques sans prendre en compte la forme sous laquelle ce contenu est présenté et la manière dont il est compris.

Le chapitre 6 se penche sur les propos de certains critiques du formalisme des positivistes logiques, qui ont cherché à définir des unités d'analyse permettant de remplacer ou de compléter la notion de théorie. La plus célèbre d'entre elles est la notion de paradigme, proposée par Kuhn (1962/1970), à laquelle je consacre la majeure partie de ce chapitre. Par une analyse critique des thèses de Kuhn, je montre dans quelle mesure l'approche que je défends en est l'héritière et en quoi elle s'en distingue. Cette analyse me permet de préciser la définition de plusieurs notions abordées dans la première partie de ce travail ; j'approfondis ainsi l'analyse de ce que j'ai appelé, au chapitre 4, la double dimension « computationnelle » et « représentationnelle » de la compréhension. Par une comparaison avec la notion kuhnienne de paradigme, je précise également le statut de l'hypothèse des versions dans ma propre démarche. Je suis finalement conduite à défendre la thèse selon laquelle l'étude de l'activité théorique gagnerait à s'intéresser aux représentations concrètes utilisées par les agents dans l'apprentissage et la pratique scientifiques, qu'il s'agisse de représentations linguistiques (comme des équations) ou de tout autre type de représentation (diagramme, schéma, etc.).

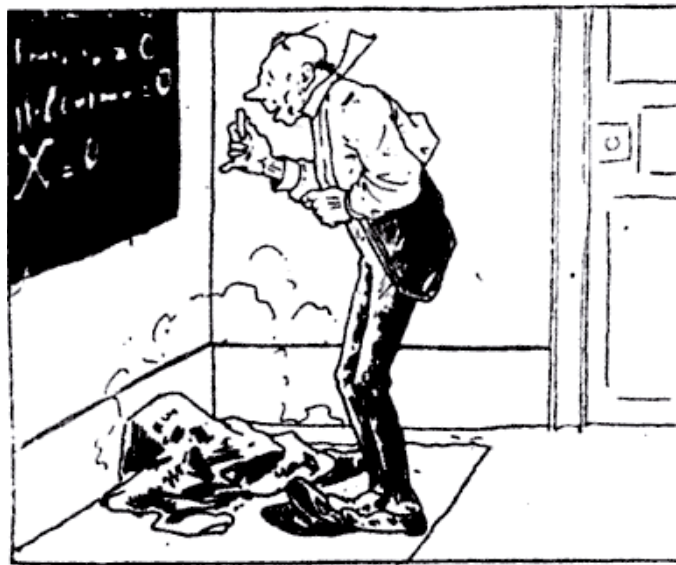
Le chapitre 7 vise à définir les outils permettant de rendre compte de la diver-

sité de ces représentations et d'analyser ses conséquences pour les raisonnements des agents. Au moyen de certains outils d'analyse des représentations définis par Goodman (1968/1976), je propose une ébauche de typologie des représentations utilisées dans la pratique scientifique, en montrant que certains types de représentations semblent plus ou moins bien adaptés à l'expression de certains contenus. Les notions que j'y définis seront au centre de mon analyse de la cartographie génétique au chapitre 9. Dans un deuxième temps, j'introduis la notion de format de représentation. Cette notion est destinée à analyser l'interaction cognitive des agents avec les représentations qu'ils utilisent, et plus particulièrement la relation entre la forme particulière de ces représentations et les capacités et intérêts cognitifs des différents agents.

La troisième partie de ce travail mettra la méthode ainsi définie à l'œuvre dans une nouvelle étude de cas.

Chapitre 5

Formuler et formaliser :
buts et limites des approches
formelles des théories scientifiques



A 3 h 1/2, le docteur découvre la valeur de x , l'inconnue cherchée. Ce qui lui cause une joie sans mélange. – Nous prions les esprits chagrins de s'abstenir de toute réflexion sur la valeur de x , et de ne point prétendre que Zéphyrin a beaucoup travaillé pour peu de chose.

L'analyse de la première partie a montré la difficulté d'isoler le contenu objectif d'une théorie pourtant canonique comme la mécanique classique sous la forme d'un petit nombre de principes joints à l'ensemble de leurs conséquences déductives. L'examen des différentes formulations sous lesquelles la mécanique est aujourd'hui enseignée et utilisée a montré que l'équivalence des principes ne garantit celle de leurs conséquences empiriques que d'un point de vue assez abstrait qui ne tient pas compte des procédures pratiques d'utilisation des différentes équations de la mécanique pour prédire et expliquer le comportement des systèmes physiques étudiés. La considération des différentes versions possibles de la mécanique, telles qu'elles ont été exprimées au cours de l'histoire par les physiciens et les philosophes qui ont cherché à en dégager la signification physique a, à son tour, mis en évidence que, selon la perspective sous laquelle on aborde les principes de la mécanique, celle-ci ne semble pas *dire* exactement la même chose du monde.

Affirmer qu'il existe un sens précis de la notion de théorie selon lequel, malgré cela, la mécanique est bien une seule et même théorie, c'est supposer qu'il est possible, au moins en principe, de définir *ce qu'elle dit* de telle sorte que ce contenu, considéré comme indépendant de la compréhension qu'en ont les agents, ne soit pas affecté par la diversité de ses formulations ; c'est donc supposer qu'*il est possible de distinguer entre ce qui relève du contenu d'une théorie et ce qui relève de sa formulation* (qu'il en existe une seule ou plusieurs, comme dans le cas de la mécanique). J'ai suggéré, au chapitre 4, que la recherche de la *meilleure* présentation des principes de la mécanique par Hertz (1894) était ainsi guidée par un idéal d'intelligibilité selon lequel comprendre parfaitement la mécanique reviendrait à se dégager de toute perspective particulière sur son contenu, et à avoir une vue globale de l'édifice, des liens logiques entre ses différents principes, et de la manière dont cet ensemble est relié au monde empirique. Tout en cherchant à approcher une présentation parfaite, Hertz semble cependant reconnaître qu'une telle entreprise est vouée à un échec au moins partiel, en raison même du fait que les théories sont des images mentales et sont, en tant que telles, déterminées par la nature particulière des esprits qui les produisent et les formulent.

C'est en partie sur le refus d'une telle conclusion que se fondent les approches des théories scientifiques, que j'appelle « formelles », à l'examen desquelles le présent chapitre est consacré. Ces approches, caractéristiques de la philosophie des sciences dominante au vingtième siècle, ont en effet cherché à développer les outils d'une *reconstruction formelle* du contenu des théories des sciences empiriques qui permette de le dégager de tous les aspects, jugés inessentiels, liés à sa (ou ses) formulation(s) particulière(s) et aux raisonnements des agents qui les développent et les utilisent. Ici, le terme de *formulation* est à comprendre en un sens plus large que celui qu'il

revêtait dans la première partie de ce travail : il sert à désigner la forme sous laquelle une théorie est habituellement présentée et utilisée par les scientifiques. En première analyse, disons qu'une formulation est l'expression du contenu d'une théorie au moyen du langage naturel augmenté des expressions mathématiques et des termes utilisés pour exprimer les concepts scientifiques qui n'ont pas toujours le sens qu'ils ont dans le langage courant (par exemple, les notions de « force » ou de « travail » en mécanique classique).

Les approches formelles reposent sur le postulat fondamental selon lequel on peut caractériser le contenu d'une théorie à l'aide d'outils purement formels, empruntés à la logique et aux mathématiques. Plus précisément, elles considèrent que *le contenu d'une théorie est ce qui est exprimé par la reconstruction formelle*. Quand il existe plusieurs formulations d'une même théorie, la reconstruction doit permettre de montrer que leur contenu est identique. C'est à ce contenu seul qu'elles s'intéressent, ignorant les aspects des formulations qui n'affectent ni ne modifient la structure logique ou mathématique de la théorie – et dont la reconstruction formelle ne rend pas compte. Pour ces approches, *ce que dit la théorie* ne dépend pas de *la manière dont elle le dit* ni de la manière dont les agents, depuis leur perspective particulière, le comprennent.

Quels sont les outils formels permettant une telle reconstruction ? C'est un des points sur lesquels se concentre le désaccord entre les deux conceptions des théories que je vais examiner dans ce chapitre, apparues successivement au vingtième siècle comme des « orthodoxies¹ » en philosophie des sciences. La première conception, baptisée « conception syntaxique » (et parfois « linguistique ») par les défenseurs de la seconde (qu'ils nomment par contraste « conception sémantique »), est celle qui émane des écrits des positivistes ou empiristes logiques, membres et héritiers du Cercle de Vienne, et en particulier Rudolf Carnap, Carl Hempel et Ernest Nagel. Son inspiration première consiste à réduire la signification de tout énoncé ou ensemble d'énoncés à son contenu empirique, en n'admettant comme éléments non empiriques que les règles de la logique. Les théories scientifiques, selon eux, doivent pouvoir être exprimées au moyen d'un langage formel, c'est-à-dire un langage tel que son vocabulaire et les règles de construction et de transformation des expressions formées au moyen de ce vocabulaire sont explicitement définis.²

Cet aspect « linguistique » de la philosophie des empiristes logiques les conduit,

¹ La première des deux a été baptisée « conception orthodoxe » [*Received View*] par Hilary Putnam (1962), un de ses critiques les plus célèbres ; la seconde, se posant en critique et remplaçante de la précédente, peut aujourd'hui, comme le suggère Soazig Le Bihan (2006), être considérée comme la « nouvelle orthodoxie ».

² Je reviens, à la fin de cette introduction, sur la notion de langage formel.

selon les défenseurs de la conception sémantique (dont les plus célèbres sont Patrick Suppes, Bas van Fraassen, Frederick Suppe et Ronald Giere), à se perdre dans des questions de philosophie du langage sans pertinence pour la philosophie des sciences. À la place de cette formalisation stricte, la conception sémantique propose l'analyse de la structure des théories scientifiques selon une méthode qui s'inspire des travaux modernes sur les fondements des mathématiques et en particulier de la théorie des modèles. Comme on va le voir, ce choix d'outils formels différents reflète des conceptions différentes du contenu des théories.

Le but même de la reconstruction formelle des théories n'est pas le même dans ces deux cadres. Comme on le verra, pour les empiristes logiques, la formalisation est un idéal³ qui fournit un critère normatif de scientificité : une théorie scientifique, pour être digne de ce nom, doit en principe pouvoir être présentée sous une certaine forme canonique. Pour la conception sémantique, le projet de reconstruction formelle se présente plutôt comme un véritable programme de recherche destiné à explorer la structure des théories scientifiques existantes. Son but est de clarifier le contenu de théories particulières et de résoudre des problèmes qui leur sont propres, aussi bien d'ordre philosophique (comme celui de la réduction d'une théorie par une autre) que d'ordre plus proprement scientifique (voir Suppes, 1968).

Le but principal de ce chapitre est de critiquer *le postulat commun* aux approches formelles, selon lequel il est possible de distinguer nettement entre ce qui relève de la formulation d'une théorie et ce qui en est indépendant. Comme il est apparu dans la première partie de ce travail, ce dernier repose, dans son ensemble, sur le postulat inverse, et vise à en tirer les conséquences ; en ce sens, l'analyse du cas de la mécanique classique, ainsi que celles des prochains chapitres, peuvent être considérées comme une *critique externe* des approches formelles : elles visent à montrer la fécondité d'une étude de l'activité théorique prêtant attention à la formulation des théories et à leur utilisation par les agents, c'est-à-dire à des aspects dont la reconstruction formelle ne rend pas compte (et ne cherche pas à rendre compte).

Dans ce chapitre, je prétends aller plus loin, et proposer une *critique interne* du projet même de reconstruction formelle des théories. En étudiant la conception du contenu des théories qui découle du présupposé opposé au mien, je souhaite non seulement, par contraste, préciser mon approche, mais encore la justifier. Par une analyse de la manière dont ces deux conceptions conçoivent la relation entre formulation et contenu, ainsi que de la critique que la conception sémantique administre à celle de l'empirisme logique, je souhaite en effet mettre en évidence des tensions internes,

³La question de la faisabilité d'une telle formalisation n'est en effet pas considérée comme un problème par les empiristes logiques.

dont je montrerai qu'elles sont partiellement dues au refus de prendre en compte les utilisateurs des théories.

Dans les sections 1 et 2, j'analyse successivement la manière dont les empiristes logiques et les défenseurs de la conception sémantique définissent le contenu des théories. Dans la section 3, j'examine la façon dont ces approches répondraient (et répondent parfois) à la question de l'identité de la mécanique classique ; cela me permettra de conclure à l'échec partiel de toute entreprise visant à distinguer nettement, au moyen d'une reconstruction formelle, entre ce qui fait partie du contenu d'une théorie et ce qui relève de sa formulation dans tel ou tel langage (section 3).

Avant d'aborder cet examen, je propose de consacrer la fin de cette introduction à une mise au point sur les notions de *langage formel*, de *formalisation* et d'*axiomatisation*, qui me seront utiles tout au long de ce chapitre.⁴

Un *langage formel* est constitué

- d'une *syntaxe* consistant en
 - un vocabulaire (un ensemble de symboles) ;
 - des règles de construction qui définissent les expressions bien formées ;
 - des règles de transformations qui indiquent les inférences permises.
- d'une *sémantique*⁵ qui est l'ensemble des règles d'*interprétation* des expressions bien formées, et permet la formation d'énoncés doués de signification (voir ci-dessous).

Formaliser une proposition ou un argument, au sens le plus strict de la notion de formalisation, consiste à les présenter sous une forme qui distingue syntaxe et sémantique. Par exemple, l'argument suivant : « Si Marie vient, Pierre vient. Or Marie vient. Donc Pierre vient » peut être formalisé de la façon suivante :

$$Pa \rightarrow Pb, Pa \vdash Pb$$

Tel quel, l'argument est purement syntaxique. La connaissance des règles de manipulation du symbole « \rightarrow » nous permet d'affirmer qu'on a là la forme d'un raisonnement valide, mais c'est la sémantique qui, nous enseignant que « a » symbolise Marie, « b » Pierre et « P » le fait de venir, nous permet d'interpréter les énoncés

⁴L'histoire de la philosophie des sciences au vingtième siècle est indissociable de celle de l'élaboration et du perfectionnement des outils formels de la logique, auxquels les philosophes des sciences de cette période, qui sont aussi, pour la plupart, des logiciens, ont contribué de manière essentielle. Dans ce chapitre, cependant, seuls les aspects indispensables à la compréhension de l'entreprise formaliste en philosophie des sciences seront mentionnés, et j'ignorerai les problèmes de philosophie de la logique qui leur sont liés.

⁵On réserve parfois l'expression « langage formel » à la seule syntaxe, considérant que la sémantique fournit une interprétation au langage formel.

qui le composent.

Une *interprétation* pour un langage formel fixe un *domaine d'objets* et y assigne des référents aux termes non logiques de son vocabulaire et une valeur de vérité aux énoncés. On appelle « modèle » d'un énoncé une interprétation dans laquelle il est vrai. On dit alors que cette interprétation « satisfait » cet énoncé. Un énoncé est vrai (par définition) de ses modèles ; il est vrai du monde si le monde en est un modèle – s'il est réalisé dans le monde. Le modèle d'un énoncé, ou d'un ensemble d'énoncés, peut être un ensemble d'éléments abstraits comme une structure mathématique, mais on peut également affirmer d'un ensemble d'objets concrets qu'il satisfait un énoncé.

Une *théorie formelle* est un ensemble d'énoncés en langage formel. Prenons un exemple géométrique simple (emprunté à van Fraassen, 1980, p. 42). Considérons la théorie T constituée des cinq énoncés suivants :

- A1 Il existe au moins une droite.
- A2 Pour deux droites quelconques, il existe au plus un point qui appartient aux deux.
- A3 Pour deux points quelconques, il existe exactement une droite qui passe par les deux.
- A4 Toute droite passe au moins par deux points.
- A5 Il existe un nombre fini de points.

Avant même de les interpréter, on peut identifier des relations logiques entre ces énoncés en vertu desquelles, par exemple, on peut affirmer que la théorie T est cohérente (il n'y a pas de contradiction logique entre ses énoncés). On peut ensuite l'interpréter, c'est-à-dire lui attribuer un domaine de référence, au sein duquel elle peut être vraie ou fausse. La figure 9 représente une structure géométrique (l'espace à sept points) qui satisfait T , et qui en est donc un modèle. Dans cette figure, sept

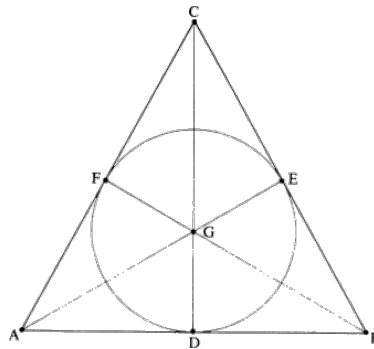


FIG. 9 – L'espace à sept points (tiré de van Fraassen, 1980, p. 42).

éléments sont appelés « points » (A, B, C, D, E, F et G), ce qui satisfait l'énoncé A5. Il est également facile de montrer que les autres énoncés sont satisfaits.

On appelle « *axiomatisation* » la présentation d'une théorie sous la forme d'un système *déductif* organisé de telle sorte que la totalité de son contenu soit comprise dans un ensemble déterminé d'énoncés qui sont les *axiomes* de la théorie. Les conséquences logiques des axiomes sont les *théorèmes* de la théorie. On appelle « clôture déductive » l'ensemble des conséquences déductives de la théorie.

À strictement parler, l'axiomatisation est indépendante de la formalisation : on peut présenter une théorie sous cette forme hypothético-déductive sans que cela implique de distinguer sa syntaxe et sa sémantique.⁶ Dans le cas d'une axiomatisation en langage formel, la théorie formelle, c'est-à-dire la structure syntaxique non interprétée de la théorie (dans l'exemple ci-dessus, les A_i , avant que leur soit attribué un domaine de discours), est aussi appelée « système formel » ou « calcul ».

On peut axiomatiser une théorie dans différents langages. Les langages se distinguent entre autres par leurs vocabulaires logique et non logique. Une formalisation *stricto sensu* (je parlerai, tout au long de ce travail, de formalisation « stricte », ou encore « syntaxique ») est faite dans un langage dont le vocabulaire non logique est *entièrement non interprété*. Les axiomatisations des théories mathématiques sont souvent faites dans un langage qui suppose déjà définies certaines notions mathématiques, celles par exemple de la théorie des ensembles, de l'algèbre matricielle, ou de toute autre branche des mathématiques. Il ne s'agit dans ce cas pas d'une formalisation au sens strict, la syntaxe n'étant pas à strictement parler non interprétée, puisqu'elle possède une interprétation mathématique préalable.

Les approches formelles des théories scientifiques, auxquelles je m'intéresse dans ce chapitre, se sont inspirées des axiomatisations en logique formelle et en mathématiques pour proposer des reconstructions formelles des théories des sciences empiriques. Dans la mesure où ces dernières, à la différence des théories mathématiques, sont supposées avoir des référents dans le monde empirique, le problème particulier qui s'y pose est celui de la mise en relation de leur structure logique ou mathématique avec les phénomènes empiriques.

C'est dans la manière dont elles abordent ce problème que réside la différence majeure entre les deux approches formelles. Alors que les empiristes logiques, comme on va le voir, *identifient l'interprétation sémantique de la théorie à ses conditions de vérité empiriques*, les défenseurs de la conception sémantique *dissocient l'interprétation de la théorie de son application empirique*. Cela se traduit entre autres par une

⁶Les premières axiomatisations (comme celle d'Euclide) n'étaient pas à strictement formalisées. Les axiomatisations modernes en logique et en mathématique supposent en revanche une formalisation préalable. L'axiomatisation de la géométrie par Hilbert, contrairement à celle d'Euclide, ne présuppose ainsi aucune connaissance de la signification de termes comme « point » et « droite », mais en donne la définition « implicite » en décrivant leurs relations.

différence dans le choix des outils formels utilisés pour axiomatiser les théories : formalisation au sens strict pour la première⁷, théorie des ensembles et autres langages mathématiques pour la seconde.⁸

Une précision s'impose, pour finir, à propos de l'usage des notions de formalisation et d'axiomatisation. Ces notions ne sont pas synonymes. Comme je l'ai dit, elles sont, à strictement parler, indépendantes l'une de l'autre.⁹ Si l'on voulait respecter le sens précis de chacune de ces expressions, on devrait parler, en ce qui concerne les approches formelles en philosophie des sciences, de formalisation uniquement pour la conception syntaxique (celle des empiristes logiques), et d'axiomatisation pour les deux conceptions des théories. Cela permettrait d'éviter certaines confusions sur lesquelles je reviendrai ; cependant, ce n'est pas toujours l'usage qu'en font les auteurs concernés : il arrive même qu'ils proposent l'usage inverse de celui que je viens de proposer.¹⁰ La plupart du temps, les termes « formalisation » et « axiomatisation » sont employés comme des synonymes ; on spécifie simplement le type d'axiomatisation / formalisation que l'on propose (formalisation stricte ou au moyen de la théorie des ensembles). Dans ce chapitre, je tâcherai de réserver le terme « formalisation » à la formalisation stricte ; pour désigner indifféremment la formalisation stricte et les axiomatisations de la conception sémantique, j'emploierai les notions d'« axiomatisation » et plus généralement de « reconstruction formelle ». Il m'arrivera cependant, par commodité, de suivre l'usage libéral des auteurs dont j'exposerai les propos, en tâchant d'éviter autant que possible les ambiguïtés. La distinction entre formalisation

⁷Les défenseurs de la conception sémantique attribuent souvent aux empiristes logiques, et en particulier à Carnap, l'ambition de formaliser les théories en logique de premier ordre. Les langages de premier ordre n'admettent comme vocabulaire logique que les quantificateurs existentiels et universels du premier ordre (ce qui veut dire qu'ils n'admettent qu'un seul type de variables, les variables d'individus) et les connecteurs propositionnels vérifonctionnels usuels. Il semble pourtant qu'une telle affirmation soit historiquement inexacte. Je reviendrai sur ce problème plus loin (section 1, note 69 et section 2, page 245).

⁸Comme on le verra à la section 2, une des raisons pour lesquelles la conception sémantique refuse le projet de formalisation en premier ordre qu'elle attribue aux empiristes logiques (voir note 7), c'est parce que les théories mathématiques indispensables à la reconstruction des théories physiques ne sont pas elles-mêmes formalisables en premier ordre. Je montrerai cependant (section 2, page 245) que ce rejet du programme des empiristes logiques repose, au delà de la question des outils disponibles, sur une conception différente du but même de la reconstruction formelle.

⁹Pour une analyse de la distinction entre formalisation, axiomatisation, symbolisation et modélisation, voir les 18 premières pages de l'article de Philippe Mongin (2003) sur l'axiomatisation en économie.

¹⁰C'est le cas, par exemple, de Frederick Suppe (1977a, p. 113) : « L'axiomatisation consiste dans l'établissement d'un calcul axiomatique, et consiste donc dans une formalisation essentiellement syntaxique [...]. La formalisation englobe à la fois les techniques syntaxiques d'axiomatisation et les techniques sémantiques de la théorie des modèles. »

au sens strict d'une part et axiomatisation ou reconstruction formelle d'autre part est en effet au cœur de mon analyse comparée des deux approches.¹¹

1 L'empirisme logique et la « signification cognitive » des théories scientifiques

Le programme formaliste en philosophie des sciences est né en Europe, dans les années 1920, avec les travaux des membres du Cercle de Vienne fondé par Moritz Schlick.¹² Ceux-ci, et tout particulièrement Rudolf Carnap, voient dans les travaux récents en logique formelle et sur les fondements de l'arithmétique de Gottlob Frege (1894), ainsi que dans les développements de l'axiomatique par David Hilbert (1899), la fondation de la théorie des ensembles par Georg Cantor (1883), et surtout le programme logiciste de Bertrand Russell et Alfred N. Whitehead (1910/1913) qui vise à réduire les mathématiques à la logique, les outils nécessaires à l'étude systématique du langage des sciences.

Situés dans ce qu'on appelle le « tournant linguistique », engagé par Wittgenstein (1922) et son célèbre slogan selon lequel « toute philosophie est “critique du langage” » (Wittgenstein, 1922, §4.0031), les membres du Cercle de Vienne ont pour projet de dénoncer les prétentions de la métaphysique en montrant que ses discours sont dépourvus de signification, ou plus précisément de ce qu'ils appellent « signification cognitive » : les discours de la métaphysique, s'ils possèdent bien une signification expressive, au même titre que la poésie ou la musique¹³, n'expriment aucune connaissance factuelle. Ils ne disent rien du monde et ne sont par conséquent pas susceptibles d'être vrais ou faux. Ce projet passe par celui d'une analyse critique du langage (Car-

¹¹Notons que dans le titre de ce chapitre, le terme « formaliser » est à comprendre au sens large de « reconstruire formellement ».

¹²Pour l'histoire du Cercle de Vienne, voir (Soulez, 1985; Bonnet et Wagner, 2006; Ouelbani, 2006). Plus généralement, pour l'histoire et l'analyse des thèses de l'empirisme logique et de Carnap en particulier, voir (Schilpp, 1963; Bouveresse, 1973; Suppe, 1977a; Jacob, 1980a,b; Friedman, 1999; Wagner, 2002a; Moulines, 2006; Chapuis-Schmitz, 2006; Bouveresse et Wagner, 2008).

¹³Les énoncés de la métaphysique « expriment le sentiment de la vie [*Lebensgefühl*] » (Carnap, 1932b, p. 175), tout comme la poésie et la musique. À la différence de ces dernières, cependant, la métaphysique « feint d'être quelque chose qu'elle n'est pas », en donnant à ses discours la forme d'une théorie, c'est-à-dire la forme d'un « système d'énoncés qui (en apparence) entretiennent mutuellement des relations de fondement » (p. 176). Carnap en conclut que, si les métaphysiciens choisissent d'exprimer le sentiment de la vie en adoptant cette forme inadéquate de discours qu'est la forme théorique, c'est qu'il leur manque « l'art d'un Beethoven pour exprimer ce sentiment dans un médium adéquat » et qu'ils sont donc « des musiciens sans talent musical. » (Carnap, 1932b, p. 177). Ces citations sont extraites de la traduction française indiquée en bibliographie (voir note 22 ci-dessous).

nap *et al.*, 1929; Carnap, 1932b), et tout particulièrement du langage de la science, qui doit donner les moyens de définir sur des bases formelles un critère précis de démarcation entre les énoncés qui sont doués d'une signification cognitive et ceux qui en sont dépourvus¹⁴ (Carnap, 1934a,b), ce qui devait permettre de tracer une frontière entre science et non science.¹⁵

L'ambition initiale du Cercle de Vienne et en particulier de Carnap consiste à réduire la signification (cognitive) de tout énoncé¹⁶ à sa signification *empirique*. Comme on va le voir, la notion de signification cognitive, dont Carnap, de ses premiers à ses derniers écrits, cherche à rendre compte par l'établissement de critères formels, repose sur le rejet, hors du champ de l'analyse philosophique du contenu des énoncés de la science, de tout ce qui a trait aux aspects psychologiques de la compréhension qu'en ont les sujets. Insistons afin de dissiper toute ambiguïté : la signification *cognitive* d'un énoncé, dans les termes de Carnap, désigne comme on va le voir sa signification objective, définie en termes de valeur de vérité ; l'emploi de l'adjectif « cognitif », ici, n'est pas à confondre avec celui qui sert à désigner aujourd'hui les processus mentaux des individus, qu'une analyse de la signification cognitive ne doit pas, selon Carnap, prendre en compte. Je me propose, dans cette section, d'examiner les conséquences d'une telle exigence sur la définition du contenu des théories des sciences empiriques.

Après avoir brièvement rappelé la motivation initiale de l'établissement d'un critère de signification cognitive (section 1.1) chez le Carnap des années 1930, je montrerai les problèmes particuliers que l'établissement d'un tel critère pose à la définition de la signification des termes et des énoncés employés dans les théories scientifiques (section 1.2). On verra ensuite (sections 1.3 et 1.4) comment ces problèmes et les débats qu'ils engendrent¹⁷ conduisent Carnap, dans les années 1950, à formuler une conception des théories scientifiques qui sera baptisée « conception orthodoxe » par Hilary Putnam (1962) puis « conception syntaxique » ou encore « linguistique » par les défenseurs de la conception sémantique. Aux sections 1.5 et 1.6, j'examinerai la version qu'Ernest Nagel (1961), généralement considéré comme un des derniers représentants de l'empirisme logique, propose de cette conception des théories ; je montrerai que Nagel se démarque plus nettement de l'empirisme logique qu'on ne

¹⁴L'analyse logique du langage de la science vise, en outre, à clarifier les fondements conceptuels et méthodologiques de la connaissance scientifique et à progresser ainsi en direction de l'unification des sciences (Carnap, 1932a).

¹⁵Sur l'assimilation de ces deux démarcations, et le désaccord entre Carnap et Popper sur ce point, voir (Jacob, 1980b, p. 131).

¹⁶Plus précisément : la signification cognitive de tout énoncé synthétique (voir plus bas).

¹⁷Ces débats se sont poursuivis principalement sur le continent américain après l'éclatement du Cercle de Vienne à la veille de la Seconde Guerre mondiale, et constituent une partie essentielle de l'histoire de la tradition que l'on appelle « positivisme logique » ou « empirisme logique ».

le reconnaît habituellement¹⁸ et prend acte, dans sa définition des théories, de l'impossibilité d'ignorer la manière dont elles sont utilisées et comprises. Cette prise de distance, jointe au maintien de certaines exigences de l'empirisme logique, conduit cependant Nagel à une position intenable, dont je montrerai qu'elle est symptomatique des limites intrinsèques du projet formaliste des empiristes logiques.

1.1 Carnap et la recherche d'un critère de signification cognitive

Intuitivement, l'exigence d'un critère de signification cognitive pour les énoncés de la science correspond à l'idée selon laquelle toute connaissance, à l'exception de la connaissance logico-mathématique, est fondée sur l'expérience. Ce qui distingue les énoncés des sciences empiriques des pseudo-énoncés de la métaphysique, c'est que les premiers présentent des états de faits et sont vrais ou faux selon que le monde est ou non tel qu'ils le décrivent.

L'établissement de ce critère repose sur la distinction entre énoncés analytiques et énoncés synthétiques. Les énoncés *analytiques* sont les énoncés qui sont vrais ou faux en vertu de leur seule forme logique ou de la signification des termes qui y figurent. Ils n'ont donc pas de contenu – ils sont *sinnlos* selon la tripartition de Wittgenstein (1922)¹⁹ – mais pas absurdes : ils ont une valeur de vérité. Les énoncés *synthétiques* sont les énoncés dont la valeur de vérité dépend des faits du monde. C'est en vertu de ces derniers que la science a un contenu cognitif, qu'elle nous dit quelque chose du monde. La définition de l'analyticité et celle du critère de signification cognitive des énoncés non analytiques ont beaucoup évolué tout au long des travaux de Carnap. Cependant, contrairement à Quine et Hempel, il ne renoncera jamais à leur possibilité.²⁰

¹⁸À l'exception, par exemple, de Rosenberg (2000) qui montre que Nagel rompt clairement avec l'empirisme logique.

¹⁹Wittgenstein distingue entre les propositions sensées, qui parlent du monde, les énoncés de la métaphysique, qui sont absurdes (*unsinnig*) et les propositions de la logique et des mathématiques, qui sont vides de contenu (*sinnlos*).

²⁰Quine (1951), dans sa critique de la distinction analytique / synthétique, distinguera entre deux sortes différentes d'énoncés analytiques : les vérités (ou faussetés) logiques, vraies (ou fausses) en vertu de leur pure forme (tautologies ou contradictions logiques, du type : « Aucun homme non marié n'est marié », où l'on peut remplacer les termes non logiques par n'importe quel terme sans affecter la valeur de vérité de l'énoncé) et les vérités (ou faussetés) analytiques qui reposent sur la signification des termes non logiques qui y figurent, et donc sur leur définition (par exemple : « Aucun célibataire n'est marié » est vrai en vertu de la signification des termes ; aucune enquête empirique n'est nécessaire pour vérifier cet énoncé, mais il faut connaître la signification de ses termes non logiques). Carnap (1952) admettra cette distinction, mais s'efforcera de maintenir la distinction entre analytique et synthétique en défendant une position conventionnaliste : la partie analytique d'un langage donné peut toujours être définie au moyen de postulats de signification, par

Comme je l'ai dit, l'ambition initiale de Carnap est de réduire entièrement la signification cognitive des énoncés non analytiques de la science à leur signification empirique.²¹ Deux énoncés (ou deux termes figurant dans des énoncés) qui ont exactement les mêmes conditions de vérité ont strictement la même signification cognitive, que leur signification expressive soit identique ou non pour certains individus :

Admettons [...] le cas où existe le critère d'un mot nouveau, comme « bébu », et qu'en effet, l'énoncé « telle chose est bébue » soit toujours vrai si et seulement si la chose est quadrangulaire. [...] Nous dirons ici : le mot « bébu » a la même signification que le mot « quadrangulaire ». Nous ne nous laisserons pas dire par ceux qui l'utilisent, qu'ils « veulent dire » autre chose, que toute chose quadrangulaire est certes aussi bébue et inversement, mais que cela tient seulement au fait que la quadrangularité est l'expression visible de la « bébitude » qui, elle, est une propriété mystérieuse, elle-même non perceptible. Nous rétorquerions qu'avec le critère se trouve également établi que « bébu » signifie « quadrangulaire », et que l'on n'est plus libre de vouloir dire ceci ou cela avec ce mot. (Carnap, 1932b, p. 159-160)²²

La *théorie vérificationniste de la signification* identifie la signification cognitive d'un énoncé synthétique à ses conditions de vérifications empiriques. Autrement dit, un énoncé synthétique a une signification cognitive si et seulement si on peut, au moins en principe, le soumettre à un test empirique, s'il est vérifiable empiriquement ; et sa signification est entièrement donnée par ses conditions de vérité. En conséquence, comprendre un énoncé, c'est en connaître les conditions de vérité et de fausseté.

Selon la version la plus stricte du vérificationnisme, un énoncé qui n'est pas directement vérifiable n'a de signification cognitive que s'il peut être réduit à un ensemble

lesquels certains termes reçoivent des définitions explicites.

²¹Jusqu'au milieu des années 1930, c'est au moyen de la seule syntaxe que Carnap entend mener l'analyse logique du langage permettant l'établissement d'un critère des énoncés doués de sens. Il voit notamment dans la théorie russellienne des types le moyen de montrer que tout énoncé dépourvu de signification est tout simplement un énoncé mal formé ; l'énoncé « César est un nombre premier » violerait ainsi les lois de la syntaxe au même titre que « César est et » (Carnap, 1934a). Il s'appuie en outre sur une distinction entre idiome formel et idiome matériel (ou contentuel) pour montrer, par des moyens syntaxiques, que les énoncés qui ont la prétention de parler de la réalité mais qui n'ont pas de signification empirique sont en fait des recommandations linguistiques déguisées (voir plus bas, note 70). Sur ce point, voir (Jacob, 1980b, p. 107). Comme le note Pierre Jacob (1980b, p. 112), après qu'Alfred Tarski (1935, 1936a,b) eut redonné « confiance aux logiciens à penchant empiriste en montrant que les concepts de la sémantique sont assis sur une base solide » Carnap (1942, §39) « se repent de n'avoir pas fait de place à la sémantique dans sa métathéorie. »

²²Toutes les citations des textes écrits par Carnap, en allemand, dans la première moitié des années 1930 (Carnap, 1932a,b, 1934a,b) sont tirées des traductions françaises indiquées en bibliographie. Les citations des textes postérieurs à 1935 (Carnap, 1939a, 1956b, 1966), écrits en anglais, sont traduites par moi-même.

fini d'énoncés protocolaires ou observationnels, c'est-à-dire directement vérifiables²³, ce qui signifie qu'il doit pouvoir être explicitement défini à l'aide d'une conjonction finie d'énoncés qui, en dernière analyse, n'emploient que des termes d'observation, et sont donc eux-mêmes directement vérifiables. Pour reprendre les termes de Hempel, l'exigence de vérifiabilité complète affirme qu'« un énoncé a une signification empirique si et seulement s'il n'est pas analytique et s'il est logiquement déductible d'une classe finie et logiquement non contradictoire d'énoncés observationnels. » (Hempel, 1965b, p. 67).

Dès le début des années 1930, Carnap abandonne cette exigence réductionniste complète, reconnaissant que de nombreux énoncés des sciences empiriques, comme par exemple les énoncés exprimant des lois de la nature, ne sont pas susceptibles d'être réduits à un nombre *fini* d'énoncés observationnels. Carnap (1936/1937) propose ainsi plusieurs critères de confirmation pour les énoncés non analytiques d'un langage, et choisit d'adopter le plus souple (un critère de confirmabilité ou de testabilité²⁴, à la place d'un critère de vérifiabilité strict). Par la suite, Carnap va continuellement libéraliser son critère de signification cognitive, admettant dès la fin des années 1930 une forme de holisme, selon laquelle certains termes et énoncés non analytiques figurant de manière essentielle dans les théories scientifiques n'ont de signification que dans le cadre de cette théorie prise comme un tout.²⁵ C'est pour rendre compte de la manière dont une théorie fournit une interprétation à ces termes et à ces énoncés que Carnap en propose la reconstruction sous la forme d'un calcul formel, et envisage l'axiomatisation des théories scientifiques²⁶ proprement dites (et non plus la seule formalisation du langage des sciences en général).

1.2 Le dilemme du théoricien

Les théories scientifiques, conçues comme des systèmes descriptifs et explicatifs des phénomènes observables, font intervenir des termes faisant référence à des choses, à des aspects des choses ou à des processus qui ne sont pas observables, comme des microparticules, des électrons, des champs (électromagnétique ou gravitationnel par

²³La question des énoncés protocolaires, devenus énoncés observationnels, a fait l'objet de nombreux débats dans les années 1930 (voir Carnap, 1932a,b; Neurath, 1933). Dès *Testability and Meaning*, Carnap (1936-1937) libéralise sa définition de l'observable, qui devient d'une importance secondaire pour la question de l'interprétation des énoncés et des termes théoriques.

²⁴Le critère de confirmabilité est le plus souple des deux ; il ne suppose pas, comme celui de testabilité, que l'on ait la maîtrise de l'expérience permettant de tester l'énoncé.

²⁵Pour une analyse critique de l'histoire du critère de signification cognitive par un empiriste logique qui, contre Carnap, y renonce, voir (Hempel, 1950, 1951, 1965b). Voir aussi (Quine, 1969).

²⁶L'axiomatisation passe ici par une formalisation au sens strict (voir page 191).

exemple). Une des étapes de la libéralisation du critère de signification cognitive par Carnap est la reconnaissance du caractère indispensable de ces termes. Dire qu'ils sont indispensables, c'est dire qu'ils ne peuvent pas recevoir de définition explicite au moyen d'énoncés ne faisant intervenir que des termes observationnels.

Afin de bien comprendre en quoi ces termes sont indispensables à l'expression des théories scientifiques, appuyons-nous sur la distinction opérée par Carnap (1956b, 1966), au sein du langage d'une théorie scientifique, entre deux sous-langages : le langage théorique (L_T) et le langage observationnel (L_O). Les énoncés de L_O emploient, outre le vocabulaire logique, commun à L_O et à L_T , et dont les termes n'ont pas de référent²⁷, uniquement des termes désignant des entités, des propriétés, des relations ou des processus observables²⁸ ; l'ensemble de ces termes constitue le vocabulaire observationnel, V_O . Les énoncés de L_O peuvent avoir une forme particulière, pour décrire un état de fait ou un événement singulier, ou une forme générale, pour exprimer une généralisation empirique ou loi phénoménale comme, par exemple, l'énoncé « à la surface de la Terre, tous les corps étendus tombent si on ne leur oppose pas de résistance ». Les énoncés de L_T sont ce que l'on appelle des lois ou principes théoriques. C'est en principe à partir d'eux que l'on peut déduire et expliquer des lois phénoménales et des descriptions empiriques particulières. Les termes qui y figurent appartiennent au vocabulaire théorique V_T et sont appelés « termes théoriques » ; ce sont les termes, mentionnés ci-dessus, qui font référence à des choses ou à des processus inobservables.

La signification des termes et des énoncés de L_O ne pose pas de problème : L_O reçoit une interprétation empirique complète. En revanche, V_T ne semble pas pouvoir être soumis à l'exigence réductionniste selon laquelle tous les énoncés de L_T devraient être déductibles d'une classe finie d'énoncés de L_O et par conséquent tout terme de V_T pouvoir être explicitement défini à l'aide d'une chaîne finie d'énoncés de L_O . En effet, exiger la réductibilité de tous les termes figurant dans une théorie à des

²⁷Pour une description plus fine des différents langages et sous-langages des théories scientifiques selon Carnap, voir la reconstruction de la conception syntaxique par Suppe (1977a), qui distingue notamment entre un langage observationnel strict, n'incluant ni quantificateur ni modalité, et un langage observationnel étendu.

²⁸Voir (Carnap, 1966, chap. 23) pour la définition de l'observable. En ce qui concerne le langage (et non les choses), la notion va progressivement se libéraliser jusqu'à devenir synonyme d'« extra-théorique » ; un terme observationnel est un terme qui est compris indépendamment de la théorie – la distinction entre ce qui est théorique et ce qui ne l'est pas dépend donc aussi de la théorie à laquelle on s'intéresse. Pour les critiques de la distinction entre langage observationnel et langage théorique, voir (Putnam, 1962; Achinstein, 1965, 1968). Hempel (1970) défend une conception des théories scientifiques généralement considérée comme proche de celle de Carnap, tout en renonçant à la distinction entre langage observationnel et langage théorique.

énoncés observationnels, c'est faire de cette théorie un simple compte rendu d'observation destiné à mettre de l'ordre dans la masse des expériences accumulées comme, pour reprendre la métaphore de Carnap, « un système de comptabilité permet d'enregistrer de façon ordonnée les dépenses et les recettes d'une société commerciale. » (Carnap, 1966, chap. 26, p. 248) Or, nous attendons d'une théorie scientifique qu'elle nous fournisse une explication en un sens plus profond que celui où une généralisation empirique comme « tous les corps incandescents sont chauds » nous fournit une explication de la chaleur qui se dégage de tel corps incandescent particulier. De plus, nous attendons d'elle qu'elle nous permette de dériver de nouvelles lois empiriques à partir de ses lois théoriques ; ces dernières ne peuvent donc pas être de simples résumés d'une conjonction de plusieurs lois empiriques. Autrement dit, exiger la réductibilité des termes théoriques au moyen d'énoncés observationnels, c'est *renoncer à une partie importante du contenu cognitif de la théorie*, qui doit nous dire quelque chose de plus que le simple résumé d'un ensemble d'observations.

Hempel (1958) nomme « dilemme du théoricien » le problème auquel conduit l'exigence réductionniste qui revient à dire que les termes théoriques doivent être entièrement définis par l'ensemble (fini) des déductions qu'ils permettent d'opérer entre différents énoncés observationnels (ce qui serait le cas si les théories étaient des manières abrégées de décrire des phénomènes observables) :

[...] si les termes et les principes généraux d'une théorie scientifique remplissent leur fonction, i.e., s'ils établissent des connexions définies entre phénomènes observables, alors on peut s'en dispenser, puisque toute chaîne de lois et d'énoncés interprétatifs établissant une telle connexion doit être remplaçable par une loi qui relie directement les antécédents observationnels aux conséquents observationnels.

En ajoutant à cette thèse cruciale deux énoncés supplémentaires dont la vérité est évidente, on obtient les prémisses d'un argument ayant la forme classique du dilemme :

- (5.1) Si les termes et principes d'une théorie remplissent leur fonction ils ne sont pas nécessaires, comme il vient d'être montré ; et s'ils ne remplissent pas leur fonction, ils ne sont évidemment pas nécessaires. Or dans toute théorie, soit ses termes et principes remplissent leur fonction, soit ils ne le font pas. Donc, les termes et principes d'aucune théorie ne sont nécessaires. (Hempel, 1958, p. 186)²⁹

La notion même de terme théorique se trouve finalement définie par le fait que ces termes ne sont pas réductibles à de l'observable : on peut dire qu'une certaine trace dans une chambre à bulles est le signe du passage d'un électron ; on peut donner

²⁹La pagination correspond à la réédition de l'article dans (Hempel, 1965b).

un ensemble de procédures destinées à mesurer la température d'un gaz. Mais la conjonction de tous les énoncés décrivant les procédures qui permettent d'affirmer, par exemple, la présence d'un électron, ne suffit pas à épuiser la signification du terme « électron ». Si cela était possible, alors tout simplement le terme « électron » cesserait d'être un terme théorique. Tant que cela n'est pas le cas, on peut espérer inventer de nouvelles procédures et découvrir de nouvelles lois empiriques à l'aide de ce terme. Le développement des théories, leur application à de nouveaux cas, la découverte de nouvelles connexions entre phénomènes, reposent sur l'existence de tels termes.

En conséquence, les règles du langage d'une théorie qui permettent la « traduction » d'énoncés de V_T en énoncés de V_O ne peuvent pas épuiser la signification des termes de V_T . À partir du moment où l'on renonce au réductionnisme, toute la question est de savoir d'où vient la partie non empirique de la signification des termes théoriques. Il s'agit d'admettre le caractère indispensable des termes théoriques, tout en garantissant la signification cognitive de la théorie.³⁰ Dès *Testability and Meaning*, Carnap (1936-1937) avait concédé que les termes théoriques ne pouvaient pas être explicitement définis en termes observationnels. Il proposait alors la notion d'énoncés de réduction bilatérale assurant la traduction du langage théorique vers le langage observationnel et vice-versa. Il va progressivement renoncer au modèle de la traduction pour adopter celui d'une *interprétation partielle*. À propos de *Testability and Meaning*, Carnap écrit ainsi en 1956 :

À l'époque de cet article, je croyais encore que tous les termes scientifiques pouvaient être introduits comme des termes dispositionnels sur la base de termes d'observation soit par des définitions explicites, soit par ce que j'appelais des énoncés de réduction, qui constituent une sorte de définitions conditionnelles [...]. Aujourd'hui, je pense, en accord avec la plupart des empiristes, que le lien entre les termes d'observation et les termes de la science théorique est bien plus indirect et faible qu'il n'était considéré dans mes écrits antérieurs et dans ceux de l'opérationnalisme. Par conséquent, un critère de signification pour L_T doit être lui aussi très faible. (Carnap, 1956b, p. 53)

Le problème est alors de définir le statut de ces règles de correspondance³¹ qui, n'étant pas des définitions explicites, ne sont pas des énoncés analytiques du langage

³⁰Le problème est en effet, comme le résume Carnap (1966, chap. 26, p. 248), de comprendre « comment on peut justifier l'usage de concepts théoriques par le scientifique sans du même coup justifier celui de termes métaphysiques par le philosophe ».

³¹Les différentes tentatives de définition de ce qui permet de relier les énoncés théoriques aux énoncés observationnels ont successivement baptisé ces règles « énoncés de réduction » (Carnap, 1936-1937), « définitions de coordination » ou de « corrélation » (Reichenbach, 1924, 1928), « définitions opérationnelles » (Bridgman, 1927, 1936), ou encore « dictionnaire » (Campbell, 1920).

de la théorie. La thèse de Carnap (1956b, 1966) est qu'une théorie prise comme un tout acquiert sa signification par le biais de règles de correspondance qui n'en donnent qu'une interprétation partielle ; les énoncés et termes théoriques tirent eux-mêmes leur signification de celle de la théorie.

1.3 Règles de correspondance et interprétation partielle (Carnap, 1956, 1966)

La volonté d'établir un critère de signification cognitive pour les énoncés de la science a progressivement conduit Carnap à proposer la reconstruction d'une théorie scientifique sous la forme d'un ensemble formalisé d'énoncés organisés en un système déductif. La notion de théorie comme calcul formel joint à une interprétation partielle apparaît dès (Carnap, 1939b) ; c'est dans (Carnap, 1956b, 1966) qu'elle reçoit son traitement le plus systématique. Les thèses exprimées dans ces textes sont celles qui correspondent au plus près à ce que les critiques des empiristes logiques ont appelé la « conception syntaxique » ou « orthodoxe » des théories scientifiques. Signalons toutefois que ces textes sont très loin, en fait, d'exprimer une orthodoxie, puisque ni Hempel, ni Nagel, pourtant considérés comme les deux autres grands représentants de cette conception, n'y adhèrent vraiment. Si Carnap est le plus orthodoxe des empiristes logiques (du moins de ceux qui écrivent encore dans les années 1950 et 1960), c'est au sens où il ne renonce jamais à certaines des inspirations initiales du Cercle de Vienne, comme l'établissement d'un critère de signification cognitive et la distinction entre énoncés analytiques et synthétiques au sein d'un langage ; corrélativement, il reste attaché à l'idée selon laquelle la formalisation d'une théorie est le moyen d'évacuer hors du champ de l'analyse philosophique tous les aspects psychologiques ayant trait à la « compréhension intuitive »³² que l'on peut avoir des termes et des énoncés des théories. Sa position n'a cependant, comme on l'a vu, jamais cessé d'évoluer.

Le contenu de cette section et de la suivante correspond aux thèses qu'il exprime en 1956 et 1966. Comme on va le voir, ces thèses comportent à la fois l'affirmation la

³² « Le développement de la physique au cours des derniers siècles, et en particulier pendant les dernières décennies, a conduit progressivement à cette méthode de construction, de test et d'application des théories physiques que l'on appelle *formalisation*, c'est-à-dire, la construction d'un calcul augmenté d'une interprétation. C'est le progrès de la connaissance et la structure particulière des domaines étudiés qui a suggéré et a rendu possible, en pratique, cette formalisation accrue. En conséquence, il est devenu de plus en plus possible de se passer d'une "compréhension intuitive" des termes abstraits et des axiomes et théorèmes formulés au moyen de ces termes. » (Carnap, 1939a, p. 67) Je remercie Delphine Chapuis-Schmitz de m'avoir indiqué ce passage, tiré d'un paragraphe intitulé « "Understanding" in Physics ».

plus explicite des présupposés de l'entreprise de reconstruction formelle des théories dans la tradition des empiristes logiques, et quelque chose comme un aveu de son échec. On verra ensuite (section 1.5) qu'Ernest Nagel (1961), tout en adoptant un certain nombre des présupposés de l'entreprise de Carnap, propose une conception des théories qui s'en distingue par des aspects importants, mettant par là en évidence les limites internes du programme empiriste logique de reconstruction formelle du contenu des théories.

Carnap (1956b, 1966) propose de concevoir une théorie scientifique comme essentiellement constituée de deux sortes de postulats : les principes théoriques ou « postulats- T », qui sont en fait les axiomes de la théorie (leur ensemble est appelé T) et les « règles de correspondance », « postulats- C » ou encore « règles- C » (leur ensemble est appelé C). Les postulats T ne contiennent que des termes de V_T , alors que les postulats C sont des énoncés mixtes, qui contiennent chacun au moins un terme théorique et un terme observationnel de manière essentielle. Leur conjonction, TC , est un système d'axiomes *partiellement interprété*. Que signifie cette notion d'interprétation partielle ?

Les principes théoriques T donnent une définition implicite des termes qui y figurent : ainsi, « $\mathbf{F} = m\mathbf{a}$ » donne une définition implicite des termes « force » et « masse » de la mécanique. À strictement parler, les termes ainsi définis n'ont pas encore de signification, pas plus que l'énoncé dans lequel ils figurent. Ce qu'on appelle ici « définition implicite » est en fait l'imposition d'une contrainte sur ce qui pourra être appelé « force » ou « masse ».³³ Cette contrainte est établie par l'expression d'une relation qui doit être satisfaite, mais aucun moyen ne nous est offert d'interpréter empiriquement ces termes et cet énoncé, qui *n'ont pas de signification extra-théorique*.³⁴ T est une théorie purement formelle, ou une forme de théorie, dont la syntaxe n'a pas encore d'interprétation sémantique.³⁵ Les termes théoriques eux-mêmes sont de purs symboles aveugles sans aucune signification empirique ; ils n'ont pour toute signifi-

³³Tout comme, dans les axiomatisations modernes de la géométrie (Hilbert, 1899), les axiomes ne donnent qu'une définition implicite de « point » et de « droite ».

³⁴Le fait que l'on utilise parfois un terme du langage ordinaire pour nommer l'entité théorique en question (comme c'est le cas, en mécanique, avec les notions de « force » ou de « travail ») ne doit pas nous induire en erreur : les représentations mentales associées au terme « force » n'en constituent pas la signification cognitive. Un terme théorique n'a pas de signification pré-théorique ou observationnelle et tire entièrement sa signification du contexte de la théorie. Notons que Hempel (1970), rejetant l'idée d'une distinction nette entre un vocabulaire théorique et un vocabulaire observationnel, refusera à l'axiomatisation des théories toute pertinence, et insistera sur l'importance des concepts « flous » et l'utilisation du langage naturel dans la formulation des théories.

³⁵Voir page 189 la mise au point sur les notions de système formel, de formalisation et d'axiomatisation.

tion que les relations que leur assigne cette structure abstraite.³⁶ Les énoncés de T ne sont pas à strictement parler des assertions, ils ne disent strictement rien : « dans la mesure où les termes théoriques de base sont seulement implicitement définis par les postulats de la théorie, les postulats n'affirment rien, puisque ce sont des formes d'énoncés plutôt que des énoncés. » (Nagel, 1961, p. 91)

Ce sont les règles de correspondance C qui fournissent son interprétation empirique au système formel T , garantissant une signification cognitive à la théorie. Elles se présentent généralement comme la spécification de procédures expérimentales pour appliquer la théorie aux phénomènes. Ces règles sont des énoncés mixtes (elles contiennent des termes de V_T et des termes de V_O) qui relient des énoncés de L_T à des énoncés de L_O . Par exemple, les énoncés « s'il se produit une oscillation électromagnétique d'une fréquence déterminée, alors on observe une certaine teinte bleu-vert », ou encore « la température (mesurée par un thermomètre et donc observable au sens large défini précédemment)³⁷ d'un gaz est proportionnelle à l'énergie cinétique moyenne de ses molécules » sont des règles de correspondance (Carnap, 1966, chap. 24, p. 233). Les règles de correspondance jettent ainsi un pont entre lois théoriques et lois empiriques (sans cela, les théories seraient parfaitement inutiles) ; c'est grâce à elles que l'on peut tirer des prédictions empiriques de la théorie.

Tout le contenu empirique de la théorie lui est fourni par les règles de correspondance. Les termes théoriques acquièrent donc une signification par le biais de l'interprétation donnée aux énoncés théoriques :

Toute l'interprétation (au sens strict du terme, i.e. interprétation observationnelle) qui peut être donnée pour L_T est donnée par les règles- C , et leur fonction est essentiellement d'interpréter certains énoncés contenant des termes descriptifs, et par là d'interpréter indirectement les termes descriptifs de V_T . (Carnap, 1956b, p. 46)

Cependant, renoncer – pour les raisons examinées – au réductionnisme en ce qui concerne les termes théoriques, c'est admettre qu'il n'existe pas de conjonction finie de règles de correspondance permettant d'épuiser la signification d'un énoncé théorique ; en outre, chaque règle de correspondance fait intervenir plusieurs termes théoriques, lesquels à leur tour figurent dans plusieurs règles de correspondance. L'*unité de signification cognitive* n'est par conséquent plus le terme ni l'énoncé, mais la *théorie*

³⁶ « [...] Quand une théorie est codifiée avec soin de telle façon qu'elle prenne la forme d'un système déductif (une tâche qui, bien que souvent difficile en pratique, est en principe réalisable), les hypothèses fondamentales de la théorie ne formulent rien d'autre qu'une structure relationnelle abstraite » (Nagel, 1961, p. 91).

³⁷ La notion d'observable fait l'objet du chapitre 23 de (Carnap, 1966) ; il en définit plusieurs sens, le sens large servant à désigner ce qui est « soit directement observable par les sens, soit mesurable au moyen de techniques relativement simples. » (p. 226)

tout entière qui, en bloc, a un ensemble de conséquences empiriques que l'on atteint grâce aux règles de correspondance.

Les règles- C établissent un lien seulement entre certains énoncés d'un type spécial de L_T et des énoncés de L_O . La thèse précédente, selon laquelle il existait des définitions de certains termes de V_T dans les termes de V_O , définitions appelées « définitions corrélatives » (Reichenbach) ou « définitions opérationnelles » (Bridgman), a été abandonnée par la plupart des empiristes comme une simplification abusive [...]. L'incomplétude essentielle de l'interprétation des termes théoriques a été mise en lumière dans mes *Fondements de la logique et des mathématiques* et est discutée par Hempel [...]³⁸. De plus, on ne peut pas exiger qu'il y ait une règle- C pour chaque terme de V_T . Si l'on a des règles- C pour certains termes, et si ces termes sont liés aux autres par les postulats de T , alors ces autres termes acquièrent eux aussi une signification observationnelle. Ce fait montre que la spécification, non seulement des règles- C , mais aussi des postulats- T , est essentielle pour le problème de la signification [*meaningfulness*]. La définition de la signification [*meaningfulness*] doit être relative à une théorie T , parce que le même terme peut être doué de signification relativement à une théorie et en être dépourvu relativement à une autre. (Carnap, 1956b, p. 48)

La thèse de l'interprétation partielle, qui repose sur l'adoption d'une position holiste, consiste à affirmer que cette interprétation empirique globale de la théorie est nécessairement incomplète : c'est ce qui rend possible l'ajout de nouvelles règles de correspondance permettant la dérivation de nouvelles lois empiriques à partir des lois théoriques (Carnap, 1966).

Un système de postulats en physique ne peut pas avoir le splendide isolement des théories mathématiques par rapport au monde. Les termes axiomatiques qu'il contient – « électron », « champ », etc. – doivent être interprétés au moyen de règles de correspondance qui les relient aux phénomènes observables. Cette interprétation est nécessairement incomplète. Parce qu'elle est toujours incomplète, le système reste ouvert, de telle sorte qu'il est possible d'y ajouter de nouvelles règles de correspondance. [...]

Il est toujours possible d'ajouter de nouvelles règles, ce qui accroît l'ensemble des interprétations disponibles pour les termes théoriques ; cependant, aussi loin qu'on l'étende, cette interprétation n'est jamais achevée. En revanche, dans un système mathématique, elle est achevée. Là, une interprétation logique d'un terme axiomatique *est* complète. (Carnap, 1966, pp. 237-238)

Le caractère incomplet de l'interprétation fournie par les règles de correspondance est donc caractéristique des *théories des sciences empiriques*, et les distinguent d'une part des théories mathématiques (« uniquement » théoriques) et d'autre part de purs comptes rendus d'observation (« uniquement » empiriques). Admettre la thèse de

³⁸Carnap mentionne (Hempel, 1951, §3) et (Hempel, 1952, §7).

l'interprétation partielle implique cependant de résoudre le problème posé par la partie de la signification de T qui n'est pas fournie par C . Voyons à présent comment Carnap prétend résoudre ce problème.

1.4 Le problème des termes théoriques et la solution de Ramsey-Carnap

Si TC ne fournit qu'une interprétation partielle des termes de V_T , en quoi consiste le reste de leur signification, leur « surplus de sens », reconnu indispensable au rôle explicatif des théories et à la possibilité de dériver de nouvelles lois empiriques à partir des lois théoriques ? Comment garantir la signification cognitive de la théorie, si une partie de sa signification, que l'on peut appeler sa signification proprement *théorique*, échappe à ses conséquences empiriques ? La thèse de l'interprétation partielle et la conception holiste de la signification des théories qui l'accompagne incitent un empiriste logique comme Hempel (1950, 1951, 1965b) à renoncer à l'idée d'un critère de signification cognitive pour les théories reposant sur une distinction nette entre la partie empirique de leur signification et ce « surplus de sens » que l'interprétation partielle ne définit pas.

En raison de ce fait [l'interprétation partielle], certains philosophes concluent que, une fois que notre critère premier a été libéralisé, nous devons admettre qu'il y a une continuité depuis les termes qui sont étroitement connectés aux observations, par exemple, « masse » et « température », *via* ceux qui en sont plus éloignés, comme « champ électromagnétique » et « fonction psi » en physique, jusqu'à ces termes qui n'ont aucun lien spécifique avec les événements observables, par exemple, ceux de la métaphysique spéculative ; par conséquent, le fait d'avoir une signification serait simplement une affaire de degré. Cette position sceptique est même adoptée par certains empiristes ; Hempel, par exemple, a donné des arguments clairs et puissants en faveur de cette conception. (Carnap, 1956b, p. 39)

Contre cette « position sceptique », Carnap va jusqu'à ses derniers travaux s'attacher à maintenir un critère de signification cognitive pour une théorie, et montrer qu'il est possible de considérer que le contenu cognitif d'une théorie peut être réduit à son contenu empirique. Tout en adoptant la thèse de l'interprétation partielle, il cherche à montrer que sa conviction ancienne selon laquelle seul le contenu proprement empirique d'une théorie compte pour la détermination de sa signification cognitive est toujours tenable.

Il s'agit donc de montrer la chose suivante : malgré le fait que l'interprétation observationnelle des termes théoriques n'épuise pas leur sens, toute question relative au « surplus de sens » de ces termes est en fait une question linguistique déguisée.

La distinction entre le contenu empirique d'une théorie, ce qu'elle nous dit sur le monde, d'une part, et les effets du langage que l'on s'est donné pour exprimer ce contenu, d'autre part, correspond à la distinction entre la partie synthétique et la partie analytique de cette théorie. C'est la raison pour laquelle Carnap affirme que, « pour la philosophie des sciences, une distinction très nette entre analytique et synthétique est [...] de la plus haute importance » (Carnap, 1966, ch. 27). Comment établir une telle distinction au sein de TC , à partir du moment où, comme on l'a vu (voir ci-dessus, page 200), les règles C ne peuvent pas être considérées comme des énoncés analytiques ?

Pour dégager la partie analytique de L_O , il suffit d'inscrire dans une théorie, en plus des vérités logiques, les postulats de signification nécessaires à rendre compte de l'usage que nous faisons des termes d'observation, comme par exemple « les baleines sont des mammifères ».³⁹ En revanche, établir de tels postulats de signification pour L_T nécessite ce qui, précisément, nous fait défaut : une connaissance de la signification de ces termes. Comment formuler des postulats de signification analogues à ceux du langage d'observation, permettant d'identifier les énoncés analytiques dans un langage théorique contenant des termes théoriques qui n'admettent une interprétation que partielle ? Carnap, par ce que Pierre Jacob (1980b, p. 155) appelle un « tour de passe-passe », connu sous le nom de méthode de « Ramsey-Carnap », propose une formalisation de la distinction entre la partie analytique et la partie synthétique des énoncés de L_T : « Je pense que, y compris pour le langage théorique, il est possible de tracer une frontière adéquate entre ce qui est scientifiquement doué de signification et ce qui ne l'est pas ».

Voici, très rapidement, en quoi consiste la solution de Carnap.⁴⁰ Elle repose sur une proposition de Ramsey (1929)⁴¹, consistant à éliminer tous les termes théoriques figurant dans l'énoncé des postulats d'une théorie TC par une manipulation logique simple : après avoir écrit TC sous la forme d'un énoncé correspondant à la conjonction de tous les énoncés qu'elle contient (l'ensemble T et l'ensemble C), on remplace tous

³⁹Pour la notion de « postulats de signification », qui permettent à Carnap de répondre à la critique quiniennne de la distinction entre analytique et synthétique, voir (Carnap, 1952, 1955). Les postulats de signification sont des conventions linguistiques données ou enregistrées à partir de l'usage en vigueur dans une communauté linguistique. Avec les règles logiques, elles constituent la partie analytique d'un langage.

⁴⁰La présentation que je fais ici de la solution de Ramsey-Carnap est très rapide et incomplète. Pour plus de détails, voir les textes de Carnap (1956b, 1966), et (Jacob, 1980b, pp. 153-155). Le travail d'Henri Galinon (2009) m'a également été d'une grande aide.

⁴¹L'article original de Ramsey, paru en 1929, est passé largement inaperçu. Pour l'histoire de la « redécouverte » de ce texte par Carnap, voir sa lettre à Hempel du 12 février 1958 (citée *in extenso* par Psillos, 1999, pp. 49-50).

les termes de V_T par des variables⁴² et on ajoute à la formule ouverte qui en résulte des quantificateurs existentiels pour chacune de ces variables. L'énoncé ainsi obtenu (la clôture existentielle de la formule), appelé « énoncé de Ramsey de la théorie » (RTC), est logiquement et empiriquement équivalent à TC . Il a en effet strictement le même contenu observationnel, mais, à la place des termes théoriques, il fait apparaître des variables liées.

Il est important de noter que les termes théoriques ont disparu de l'énoncé de Ramsey. Ils sont remplacés par des variables. La variable « C_1 » ne réfère à aucune classe particulière. Tout ce qui est affirmé, c'est qu'il existe au moins une classe qui satisfait certaines conditions. La signification de l'énoncé de Ramsey n'est en aucune façon modifiée par un changement arbitraire de variables. Par exemple, les symboles « C_1 » et « C_2 » peuvent être intervertis ou remplacés par d'autres variables arbitrairement choisies, telles que « X_1 » et « X_2 ». La signification de l'énoncé reste la même. (Carnap, 1966, chap. 26, p. 251)

RTC exprime donc la partie synthétique de TC . Tout ce que dit TC à propos du monde est contenu dans RTC . Reste à présent à rendre compte de la partie analytique de TC . Carnap propose de concevoir TC comme équivalente à la conjonction de deux énoncés, A_T , vide de tout contenu factuel et qui joue le rôle des postulats de signification pour tous les termes de V_T , et F_T , qui énonce tout le contenu factuel de la théorie (c'est donc RTC). On a donc :

$$(A_T \wedge RTC) \vdash TC$$

Le postulat A_T d'une théorie TC , qui n'est pas affirmé par la théorie, est :

$$RTC \rightarrow TC.$$

Cet énoncé est empiriquement vide⁴³ : tout le contenu factuel de TC est dans RTC . L'énoncé A_T épuise bien tout ce que la théorie dit des termes de V_T : il dit simplement que si RTC est vrai (c'est-à-dire si la théorie est réalisée, si le monde empirique en est un modèle), alors nous devons comprendre les termes théoriques de telle façon que la théorie entière soit vraie. C'est un énoncé analytique, puisque sa vérité ne dépend que de la signification attribuée aux termes théoriques.

La théorie TC est impliquée par la conjonction de RTC et du postulat A_T exprimé par l'implication « $RTC \rightarrow TC$ » :

$$(RTC \wedge (RTC \rightarrow TC)) \vdash TC.$$

Cette « solution » permet notamment à Carnap, comme on le verra à la section 1.6.3, d'affirmer que le débat entre réalisme et instrumentalisme se réduit à une

⁴²Plus précisément : on remplace tous les termes théoriques par des noms de classes et de relations : par exemple, on remplace toutes les occurrences de « molécule » ou « électron » par une variable x ou y , dont on précisera qu'elle appartient à une certaine classe (« Mol », ou « Elec »).

⁴³Son énoncé de Ramsey est une tautologie (merci à Henri Galinon pour cette remarque).

question pragmatique de choix de langage. Il montre en effet que la question de l'existence des entités auxquels les termes théoriques ont l'air de référer est une question linguistique déguisée.

Au moyen de cette manipulation logique, Carnap semble donc être parvenu à rendre « inoffensif » le surplus de sens des termes théoriques. Cependant, affirmer que *RTC* et *TC* disent exactement la même chose *à propos du monde* ne suffit pas, comme le reconnaît Carnap, à affirmer qu'ils disent exactement la même chose *tout court* :

Ramsey a seulement voulu mettre en évidence la *possibilité* de formuler n'importe quelle théorie dans un langage qui n'exige pas de termes théoriques mais qui dit cependant la même chose que le langage habituellement utilisé par les scientifiques.

Quand nous disons qu'il « dit la même chose », nous ne prenons en compte que les conséquences observables. Nous ne prétendons pas, bien sûr, que le langage de Ramsey dise *exactement* la même chose. [Le langage habituellement utilisé par les scientifiques] présuppose que des termes théoriques comme « électron » ou « masse » désignent quelque chose *de plus* que ce qui nous est fourni par le contexte de la théorie elle-même. C'est ce que certains auteurs ont appelé le « surplus de sens » d'un terme. Quand on prend en compte ce surplus de sens, les deux langages ne sont certainement plus équivalents. L'énoncé de Ramsey représente le *contenu d'observation* tout entier de la théorie. Le grand mérite de Ramsey est d'avoir compris que *la théorie ne requiert que ce contenu observationnel pour remplir sa fonction de théorie, c'est-à-dire pour expliquer les faits connus et en prédire de nouveaux*.⁴⁴ (Carnap, 1966, chap. 26, p. 254)

Carnap formule ici quelque chose qui peut ressembler à un aveu d'échec, suivi immédiatement de l'affirmation la plus explicite qui soit du postulat fondamental de son entreprise formaliste. En effet, Carnap reconnaît, d'une part, qu'il existe un « surplus de sens » des termes théoriques, laissé de côté par l'interprétation partielle que fournissent les règles de correspondance. Dans la mesure où ce surplus de sens ne fait pas partie du contenu empirique de la théorie, il semble qu'une partie de sa signification soit irréductible à sa signification empirique. D'autre part, il évacue ce problème par ce que l'on peut décrire comme un véritable coup de force : ce surplus de sens n'a pas d'importance car il ne joue aucun rôle, affirme Carnap, dans la fonction prédictive et explicative de la théorie. Ce qui lui permet d'affirmer cela, c'est l'affirmation que cette fonction prédictive et explicative de la théorie repose tout entière sur son contenu observationnel : « l'énoncé de Ramsey possède exactement le même *pouvoir explicatif et prédictif* que le système de postulats d'origine. » (Carnap, 1966, chap. 26, p. 252) Autrement dit, si l'on s'intéresse à la théorie *en tant qu'elle permet de fournir*

⁴⁴ Je souligne.

des prédictions et des explications, alors son contenu peut, selon Carnap, se réduire à son contenu empirique. Plus qu'une solution au problème, Carnap explicite ici le présupposé fondamental que son entreprise doit assumer, et dont je vais montrer qu'il n'est pas tenable.

Comme on va le voir, la conception des théories proposée par Ernest Nagel (1961), telle que je la comprends, repose sur le *refus du présupposé selon lequel la fonction prédictive et explicative d'une théorie est entièrement réductible à l'expression de son contenu observationnel*. Cela me permettra de montrer que les différentes manières de comprendre et de rendre compte de la thèse – ou, plutôt, du problème – de l'interprétation partielle impliquent autant de conceptions différentes de la signification des théories et de la notion de théorie.⁴⁵ Leur examen me conduira à montrer le caractère intenable du projet formaliste des empiristes logiques.⁴⁶

1.5 Ernest Nagel (1961) et l'importance des modèles

Dans les chapitres 5 et 6 (respectivement intitulés « Experimental Laws and Theories » et « The Cognitive Status of Theories ») de *The Structure of Science*, Ernest Nagel (1961), considéré comme l'un des derniers représentants de l'empirisme logique, présente sa conception des théories scientifiques. Contemporaine des dernières versions de celle de Carnap (1956b, 1966) que je viens de présenter, la conception de Nagel en partage certains aspects importants.⁴⁷ Elle s'en distingue cependant sur des points essentiels, en particulier sur la question de l'interprétation du système formel de la théorie (*T*).

Comme on va le voir, les difficultés soulevées par la position de Nagel permettent de saisir les limites de la reconstruction formelle des théories proposée par les empiristes logiques. Il y introduit en effet une troisième composante (en plus des deux – *T* et *C* – définies par Carnap), dont on va voir qu'elle en constitue une modification majeure, justifiée par les limites de la thèse de l'interprétation partielle. Prendre au sérieux la nécessité de considérer cette troisième composante, dont je vais montrer qu'elle correspond à la prise en compte de l'utilisation et de la compréhension des théories par les agents, implique selon moi de renoncer au programme formaliste des

⁴⁵ La notion d'interprétation partielle a fait l'objet de nombreuses critiques (voir notamment Putnam, 1962; Achinstein, 1963) montrant qu'il est difficile de lui donner un sens clair et satisfaisant; ayant été introduite sans définition dans (Carnap, 1939b), elle est « utilisée par Carnap et d'autres [...] au moyen d'abondantes références croisées, mais sans plus d'explication. » (Putnam, 1962, p. 245)

⁴⁶ Il permettra également de comprendre pourquoi Suppe (1971) a vu dans la thèse de l'interprétation partielle une préfiguration de la conception sémantique des théories.

⁴⁷ Carnap (1966) y fait d'ailleurs plusieurs références explicites qui semblent manifester un accord profond.

empiristes logiques. En restant cependant paradoxalement attaché à ce programme, Nagel aboutit à une conception intenable de la signification des théories. Tenter de mettre de l'ordre et de clarifier les difficultés et parfois les contradictions apparentes de ce texte dense me permettra de montrer qu'un certain nombre de problèmes qui sont les axes majeurs de l'approche que je défends sont responsables de certaines tensions internes du projet des empiristes logiques. Le texte de Nagel est en effet riche de suggestions qui ont fortement inspiré mon travail.

1.5.1 Les trois composantes de la théorie

Nagel (1961) propose de concevoir les théories scientifiques comme essentiellement constituées de trois composantes (et non pas deux, comme chez Carnap) :

- (1) un système formel abstrait qui est le squelette logique du système explicatif, et qui « définit implicitement » les notions de base du système ; (2) un ensemble de règles qui assignent un contenu empirique au système formel abstrait en le reliant aux objets concrets de l'observation et de l'expérience ; et (3) une interprétation ou modèle pour le système formel abstrait, qui fournit de la chair à la structure du squelette à l'aide d'une étoffe conceptuelle ou visualisable plus ou moins familière (Nagel, 1961, p. 90).

On voit que les deux premières composantes (squelette logique et règles de correspondance) correspondent conjointement à *TC*. En ce qui les concerne, Nagel ne se démarque pas de la position de Carnap exposée précédemment. Les règles de correspondance ont pour fonction de relier les axiomes du squelette aux phénomènes empiriques ; elles sont absolument indispensables.⁴⁸

À quoi sert la troisième composante introduite par Nagel ? Nagel nous dit que le « modèle » est ce qui fournit une interprétation (il parle d'ailleurs indifféremment du « modèle » ou de l'« interprétation ») aux axiomes qui, sans lui, forment une structure purement syntaxique sans signification (un squelette sans « chair »). Les règles de correspondance ne sont-elles pourtant pas déjà chargées de fournir cette interprétation (partielle) aux axiomes ? Rappelons en effet que, pour Carnap, l'interprétation de L_T « au sens strict du terme » (Carnap, 1956b, p. 46), et donc la seule qui doive entrer en compte pour étudier la signification cognitive de la théorie, est l'interprétation observationnelle que lui fournissent les règles de correspondance. Qu'entend donc Nagel par « interprétation » pour affirmer qu'une troisième composante est nécessaire à l'interprétation des axiomes de la théorie, tout en reconnaissant

⁴⁸ « Le caractère indispensable de ces liaisons a été constamment souligné dans la littérature récente, et une diversité de noms leur ont été attribués : définitions de coordination, définitions opérationnelles, règles sémantiques, corrélations épistémiques et règles d'interprétation. » (Nagel, 1961, p. 93)

le caractère indispensable des règles de correspondance ?

Le problème réside dans la notion d'interprétation partielle, dont j'ai signalé précédemment le caractère mal défini. Plusieurs auteurs, comme Achinstein (1963), affirment que, dans la mesure où les règles de correspondance ne sont pas de simples postulats de signification des termes de V_T , et donc dans la mesure où C n'est pas strictement analytique, ces règles non seulement ne donnent pas une interprétation complète du système formel T , mais même ne lui fournissent aucune interprétation.⁴⁹ Autrement dit, à partir du moment où les règles de correspondance ne fournissent pas une interprétation complète au système formel, il faut, selon lui, renoncer entièrement à l'idée qu'elles lui fournissent une interprétation, même partielle. Il semble que cela soit également la lecture de Nagel, qui suggère que les règles de correspondance ne suffisent pas à transformer les axiomes non interprétés du système formel en énoncés doués de signification. Comme l'examen de la section 1.6 le fera mieux comprendre, ces règles sont plutôt, pour Nagel, des « recettes » permettant, à partir du squelette syntaxique, de tirer des prédictions empiriques, mais elles ne *font pas dire* quoique ce soit aux énoncés de ce squelette. Qu'il soit possible ou non de « sauver » la notion d'interprétation partielle en lui donnant tel ou tel sens⁵⁰ n'est pas précisément la question à laquelle je cherche à répondre. Ce qui m'intéresse ici est que, dans le refus de Nagel de voir dans les règles de correspondance un moyen de fournir une interprétation au système formel, transparaît une conception différente du contenu des théories scientifiques, et de la notion de théorie elle-même. Elle se manifestera de manière encore plus évidente dans la discussion de la section 1.6. Les deux questions qui se poseront alors seront de savoir, premièrement, si l'infléchissement de la notion de théorie proposé par Nagel est nécessaire et, deuxièmement, si une telle conception est compatible avec l'entreprise formelle.

Pour l'heure, admettons, avec Nagel, que l'interprétation de la théorie ne lui est pas fournie par les règles de correspondance et donc que la notion d'interprétation partielle ne suffit pas à rendre compte de son contenu cognitif. C'est là une différence majeure avec la conception de Carnap : selon Nagel, les conséquences empiriques de la théorie n'en constituent pas tout le contenu cognitif. Cela signifie que l'ensemble constitué par T et C , à supposer qu'il puisse servir à tirer des prédictions, *ne dit rien*. Si la théorie se réduit à cela, elle remplit peut-être son rôle prédictif, mais ce rôle se réduit à celui d'une machine à calculer : on peut (en principe), en manipulant

⁴⁹Cependant, Suppe (1971) affirme que l'argument d'Achinstein repose sur une mauvaise compréhension de la notion d'interprétation partielle telle que Carnap (1956b) la présente.

⁵⁰C'est ce que fait Suppe (1971), qui y voit, en germe, la conception sémantique ; sa reconstruction de la notion repose cependant sur l'abandon de la distinction entre langage théorique et langage observationnel. Elle ne peut donc pas, comme il le reconnaît, être attribuée à Carnap.

cet ensemble de symboles suivant certaines règles, tirer des prédictions empiriques ; mais la théorie, ainsi définie, est une « boîte noire » qui, en tant que telle, ne nous dit strictement rien. Afin de lui faire dire quelque chose, il faut lui ajouter une troisième composante, que Nagel appelle son (ou ses) modèle(s).

En quoi le modèle de la théorie doit-il donc consister pour faire de T un ensemble d'énoncés doués de sens ? Le terme « modèle » présente l'inconvénient d'une très grande polysémie, que les propos de Nagel ne contribuent pas à réduire. On peut en première analyse dégager deux sens différents de la notion telle qu'elle apparaît dans ces pages.

En premier lieu, Nagel semble avoir à l'esprit (notamment page 96 de son ouvrage) la notion logique de modèle (voir ci-dessus, page 190) : les modèles sont des structures qui fournissent une interprétation à une théorie formelle et la satisfont. Cela signifie que les relations entre les éléments de la structure sont adéquatement décrites par la théorie. Un modèle en ce sens peut être une structure abstraite (mathématique), mais également un ensemble d'éléments concrets quiinstancient certaines relations.

Deuxièmement, Nagel emploie également le terme « modèle » pour désigner un système qui, par certains de ses aspects, entretient une relation d'isomorphisme, ou plus largement d'analogie, avec un autre système, dont il sert par conséquent de représentation. On parle alors de modèle analogique, ou encore iconique.⁵¹ Ainsi, le modèle miniature d'un avion, en raison d'une certaine identité de structure avec l'avion en taille réelle, en est une représentation. Ainsi, la théorie cinétique des gaz utilise le modèle des boules de billard pour représenter le mouvement des molécules : elle représente ce mouvement comme analogue au mouvement observable et bien connu des corps macroscopiques ; elle nous invite à nous représenter les molécules du gaz *comme* des boules de billard. Ces analogies, que Nagel appelle « matérielles » (*substantial*), servent la plupart du temps à se représenter des phénomènes inobservables (parce que, par exemple, microscopiques) et mal connus sous les traits de phénomènes et d'objets familiers.

[...] quand les physiciens parlent du modèle d'une théorie, ils ont presque toujours en tête un système d'objets dont la taille diffère très largement des objets qui

⁵¹La notion de modèle iconique est déjà discutée par Campbell (1920). Dans les années 1960, plusieurs philosophes (en particulier Hesse, 1966; Black, 1962) la développent et soulignent l'importance des analogies et des métaphores dans les théories scientifiques. Hesse insiste en particulier sur le rôle des modèles analogiques dans l'explication. La notion de modèle et le rôle des analogies et des métaphores dans le raisonnement scientifique fait aujourd'hui l'objet d'un certain nombre de travaux en philosophie des sciences, qui placent l'utilisation de modèles au cœur de l'activité théorique (voir par exemple Magnani et Nersessian, 2002; Magnani *et al.*, 1999). Ces dernières approches sont en rupture complète avec celle des empiristes logiques. Ici, ce qui m'intéresse est la manière dont Nagel cherche à intégrer la notion de modèle à reconstruction formelle dans la tradition empiriste logique.

sont, au moins approximativement, réalisables dans l'expérience familière, de telle façon qu'un modèle dans ce sens peut être représenté picturalement ou en imagination. (Nagel, 1961, p. 110)

Aux analogies matérielles, Nagel ajoute un second type d'analogies, qu'il appelle « formelles ». Ce sont celles suggérées par l'usage, dans un certain domaine scientifique, d'un formalisme (par exemple un type d'équations) appartenant à un autre domaine. Ce qui est analogue entre les deux domaines est alors une structure de relations abstraites. Certains principes de la théorie de la relativité, par exemple, sont formulés par analogie avec le principe fondamental de la dynamique newtonienne : les équations qui les expriment sont du même type mathématique : « Cet exemple montre comment le formalisme mathématique d'une théorie peut servir de modèle pour la construction d'une autre théorie au domaine d'application plus large que celui de la théorie d'origine. » (Nagel, 1961, p. 111)

Remarquons dès à présent que la différence entre les deux types d'analogie est moins marquée qu'il n'y peut paraître. En effet, analogies formelle et matérielle vont souvent de pair. Prenons par exemple la théorie cinétique des gaz : le modèle (matériel) des boules de billard est suggéré par le fait que cette théorie utilise des équations du type de celles utilisées pour décrire le mouvement des corps macroscopiques (analogie formelle). Quand on abstrait du système concret des boules de billard ses aspects non pertinents pour l'étude du mouvement des molécules (couleur, numéro, masse, composition chimique, par exemple), l'analogie qui demeure est bien une analogie structurelle dans le comportement mécanique, analogie incarnée par le type d'équations utilisées, qui « suggèrent » donc l'image de systèmes mécaniques familiers comme les boules de billard pour illustrer la théorie.

Inversement, l'importance heuristique des analogies matérielles pour le développement de l'appareil mathématique (et donc des analogies formelles) d'une théorie est largement reconnue. Le rôle de la visualisation, de l'utilisation d'images et d'analogies matérielles pour l'élaboration et l'exploration d'hypothèses, leur mise en forme mathématique, ainsi que l'extension de leur domaine d'application, a fait l'objet de plusieurs témoignages et réflexions philosophiques ou méta-scientifiques de la part des physiciens eux-mêmes. Maxwell (1855), qui développe une véritable méthodologie de l'analogie physique (voir Nersessian, 2002b), en est un exemple fameux : il explique en particulier comment l'analogie visuelle, purement qualitative, proposée par Faraday pour représenter le champ électromagnétique sous la forme de « lignes de forces » du type de celles que dessine la limaille de fer en présence d'une source magnétique, l'a conduit, par abstraction, à formuler sa théorie mathématique de

l'électromagnétisme.⁵²

On comprend en quel sens Nagel affirme que les modèles donnent une « chair » au squelette syntaxique de la théorie. Cependant, il n'est pas encore évident de voir l'unité qu'il suggère entre les modèles comme interprétation d'un squelette syntaxique, les modèles matériels comme analogie visuelle et les modèles portés par l'usage de certaines formes d'équations. Cette apparente hétérogénéité est symptomatique des difficultés qui m'intéressent ici ; elle devrait s'éclaircir par la suite. Admettons pour l'instant, en un sens encore métaphorique, que les modèles sont ce qui permet de *se représenter*, de *se figurer*, une théorie autrement inintelligible.

1.5.2 Les modèles et l'intelligibilité de la théorie : explication des phénomènes et compréhension de la théorie

Outre l'importance heuristique des modèles, Nagel affirme que ces derniers jouent un rôle fondamental dans l'explication. Avant d'analyser ses propos, il convient de dire quelques mots de la conception orthodoxe de l'explication scientifique dite « déductive-nomologique » de Hempel⁵³, que Nagel, par ailleurs, contribue à développer.⁵⁴ Selon ce modèle, une explication a exactement la même forme logique qu'une prédiction : l'explication scientifique d'un phénomène particulier ou d'une régularité phénoménale a la forme d'une déduction de type *modus ponens* dont les prémisses doivent contenir de manière essentielle une loi universelle (d'où l'adjectif « nomologique »). Ce qui différencie prédiction et explication – outre le fait que la conclusion de la déduction, dans le cas de la prédiction, décrit un phénomène qui ne s'est pas encore produit – est d'ordre psychologique et « pragmatique » ; c'est le « sentiment de compréhension » qui naît d'une bonne explication. De même que, pour Carnap, les phénomènes psychologiques liés aux « représentations » ou « images » suggérées par les mots employés ne sont d'aucune pertinence pour la détermination de la signification cognitive d'un énoncé, les aspects psychologiques de la compréhension n'ont pas

⁵²L'utilisation des lignes de forces de Faraday par Maxwell, et plus largement les réflexions méthodologiques de ce dernier sur l'utilité heuristique d'analogies et de métaphores a fait l'objet de nombreuses études de cas dans la littérature récente en philosophie des sciences sur le changement conceptuel et le raisonnement sur la base de modèles [*model-based reasoning*] (voir notamment Nersessian, 1984, 1992a, 2002b; Cat, 2001).

⁵³Le texte fondateur de la conception hempélienne de l'explication est (Hempel et Oppenheim, 1948). Hempel développera sa conception dans de nombreux articles, en réponse aux critiques qui lui sont adressées (voir notamment Hempel, 1965a). Pour une analyse détaillée de l'histoire des débats sur l'explication, voir (Salmon, 1989; Salmon et Kitcher, 1989).

⁵⁴Le sous-titre de *The Structure of Science* est « Problems in the Logic of Scientific Explanation ». L'ouvrage se présente en majeure partie comme un développement de la conception empiriste logique de l'explication scientifique.

à être pris en compte, dans l'orthodoxie hempélienne, par l'analyse philosophique de l'explication scientifique : « des expressions comme “domaine de la compréhension” et “compréhensible” n'appartiennent pas au vocabulaire de la logique, car elles font référence aux aspects psychologiques et pragmatiques de l'explication. » (Hempel, 1965a, p. 413).

Au sens très large, expliquer quelque chose à quelqu'un est rendre cette chose claire et intelligible pour cette personne, c'est-à-dire la lui faire comprendre. Ainsi conçu, le mot « explication » et ses cousins sont des termes pragmatiques : leur usage nécessite de faire référence aux personnes impliquées dans le processus de l'explication. [...] L'explication en ce sens pragmatique est donc une notion relative : quelque chose peut être dit de manière significative constituer une explication en ce sens uniquement pour tel ou tel individu. (Hempel, 1965a, pp. 425-426)

On a vu dans ce qui précède que tout le contenu empirique de la théorie est fourni par le squelette syntaxique et les règles de correspondance, qui permettent en principe de tirer des prédictions empiriques à partir des axiomes. Si l'on affirme comme Hempel que l'explication a la même forme logique que la prédiction, on voit mal comment les modèles pourraient ajouter quoi que ce soit à la fonction explicative des théories.

C'est donc une notion d'explication qui s'éloigne de la définition orthodoxe que Nagel semble avoir à l'esprit quand il affirme que, dans la mesure où les axiomes du squelette ne sont que des formes d'énoncés qui ne disent rien sur le monde, et parce que les règles de correspondance n'en fournissent pas d'interprétation (ne les transforment pas en énoncés proprement dits), le fait que l'on puisse en tirer des énoncés décrivant les phénomènes observables ne constitue pas une explication. Ici, la notion d'explication a manifestement le sens que Hempel écarte dans les propos rapportés ci-dessus ; au sens où Nagel l'emploie ici, on doit pouvoir attendre d'une explication qu'elle nous permette de comprendre les phénomènes.⁵⁵ La prédiction (ou l'explication selon le modèle DN) qu'on en tire est un pur calcul non compris, que pourrait faire un ordinateur. Réduite à ses deux premières composantes, la théorie ne nous permet pas de comprendre les phénomènes, car elle n'est pas elle-même *intelligible* : sans un modèle, on n'a aucun moyen de comprendre ce qu'elle dit.

⁵⁵Cette rupture de Nagel avec les dogmes de l'empirisme logique a été soulignée par Alexander Rosenberg (2000). Selon lui, le naturalisme aujourd'hui est un « mouvement en philosophie qui repose sur l'adoption des arguments de Quine contre l'empirisme, sans pour autant accompagner Quine jusqu'à l'éliminativisme concernant les modalités ou la signification. C'est la raison pour laquelle les naturalistes sont les enfants d'Ernest Nagel. Le *locus classicus* du naturalisme philosophique n'est pas “l'épistémologie naturalisée” (Quine, 1969). C'est bien plutôt *The Structure of Science* (Nagel, 1961). » (Rosenberg, 2000, p. 7)

Nagel remarque que l'histoire des sciences est riche de périodes de doute où une théorie ne parvient pas à s'imposer en tant que telle parce que, précisément, aucun modèle intelligible n'en est disponible. Ainsi, la difficulté de trouver un modèle « visualisable » de la mécanique quantique (voir ci-dessus, chapitre 4, section 2.1) a longtemps été un obstacle pour considérer qu'elle donnait une explication satisfaisante des phénomènes.⁵⁶ Affirmer qu'il n'existe pas de modèle « visualisable » de la mécanique quantique tant que l'on ne peut pas en présenter un modèle classique, c'est dire qu'il nous est impossible de *nous représenter* le mouvement des corps microscopiques autrement que comme celui, bien connu et décrit par la théorie classique, des corps à l'échelle macroscopique. Donner un modèle mécanique de phénomènes encore mal compris, permettant par là de dresser une analogie (matérielle ou formelle) avec des phénomènes classiques, est une manière de rendre les nouveaux phénomènes intelligibles. Comme je l'ai mentionné au chapitre 4 (section 2.1), plusieurs physiciens de la fin du dix-neuvième et du début du vingtième siècles, comme Maxwell, Kelvin ou Boltzmann, affirment ainsi qu'un phénomène n'est compris que quand la théorie qui l'explique peut en fournir un modèle mécanique :

La passion de Lord Kelvin pour les modèles mécaniques est un exemple fameux d'une telle attitude; il ne s'est jamais senti parfaitement à l'aise avec la théorie électromagnétique de la lumière de Maxwell parce qu'il était incapable d'en figurer un modèle mécanique satisfaisant. (Nagel, 1961, p. 114)

Utiliser un formalisme dont on est familier (comme celui de la mécanique classique) pour décrire des phénomènes mal connus est une tentative de les représenter comme des phénomènes classiques. On peut ainsi appeler « modèle mécanique » une représentation imagée comme celle des boules de billard, mais également l'utilisation d'équations de la forme $\mathbf{F} = m\mathbf{a}$. En ce sens, utiliser un certain langage mathématique, un certain type d'équations, c'est déjà présenter le squelette syntaxique de la théorie (en fait inaccessible) sous un certain habillage. C'est là un point important, sur lequel Duhem (1914) était explicite (voir ci-dessus, chapitre 4, section 1.1, pages 164-164) : l'utilisation de l'algèbre symbolique, que Duhem appelle « calcul » est, en tant que telle, déjà une manière d'« habiller » la théorie définie comme un ensemble de symboles abstraits. Alors que l'intelligence parfaite de la théorie (celle des esprits abstraits) consiste en la manipulation de ces symboles abstraits au moyen du seul raisonnement logique, les esprits amples, comme on l'a vu, « calculent » : « cette manœuvre des symboles algébriques qu'on peut, dans la plus large acception du mot,

⁵⁶Pour une analyse de la manière dont les fondateurs de la mécanique quantique, et en particulier Niels Bohr (1961), avec la notion de « complémentarité », ont cherché à s'accommoder de l'impossibilité de présenter la mécanique quantique dans un cadre classique en garantissant l'intelligibilité, voir (Chevalley, 1991). Voir aussi les autres références indiquées au chapitre 4, note 14.

nommer le *calcul*, suppose, chez celui qui la crée comme chez celui qui l'emploie, bien moins la puissance d'abstraire et l'habileté à conduire par ordre ses pensées, que l'aptitude à se représenter les combinaisons diverses et compliquées qui se peuvent former avec certains signes visibles et dessinables » (Duhem, 1914, p. 90). En ce sens, le langage mathématique *n'est pas* une syntaxe non interprétée.

Le modèle est à la fois ce qui permet de se représenter des phénomènes mal connus *comme* des phénomènes familiers, et de se représenter, de se *figurer* la théorie elle-même. Le modèle des boules de billard est à la fois *une représentation des molécules* et *une visualisation de ce que dit la théorie*. Comprendre les phénomènes à l'aide d'une théorie implique que l'on soit en mesure de comprendre la théorie elle-même, en tant qu'elle dit quelque chose, qu'elle a un contenu cognitif. Or ce contenu, c'est le modèle qui nous permet de nous le figurer. C'est en ce sens que le modèle fournit une interprétation du squelette syntaxique. Les différents sens de la notion de modèle dégagés précédemment commencent ainsi à trouver une unité (sans que le statut de cette « troisième composante » soit encore tout à fait clair).

Cette analyse repose sur une notion de compréhension qui s'éloigne fortement des exigences carnapiennes : Nagel souligne en effet que cet « habillage » des théories dépend du contexte de leur utilisation : un type d'équations duquel on est familier, comme celles qui incarnent un modèle mécanique, devient, en raison même de cette familiarité avec les théories qui l'utilisent, un standard d'intelligibilité. Or, les standards d'intelligibilité sont voués à changer.⁵⁷ Nagel évoque ainsi des « modes dans les préférences que les scientifiques manifestent pour différents types de modèles, de type matériel aussi bien que formel » (Nagel, 1961, p. 115). La « passion » de Lord Kelvin pour les modèles mécaniques illustre ce point. L'autre exemple souvent pris par Nagel est celui de la mécanique quantique ; le fait qu'elle puisse sembler inintelligible (parce qu'on ne peut, comme on l'a vu, en présenter un modèle classique) n'implique pas que ce sera toujours le cas :

Le fait qu'aucun modèle visualisable incarnant les lois de la physique classique ne puisse être donné de la théorie quantique, [objectent les défenseurs de la mécanique quantique], n'est pas une base suffisante pour nier que la théorie quantique décrive les propriétés structurelles des processus subatomiques. Il est indéniablement souhaitable d'avoir un modèle satisfaisant de la théorie. Mais

⁵⁷ Voir ci-dessus, chapitre 4 (section 2.1), ma présentation de la conception évolutionniste des lois de la pensée défendue par Boltzmann : comme je l'ai souligné, pour Boltzmann, si les modèles mécaniques sont seuls, à son époque, susceptibles de fournir une représentation intelligible des phénomènes – et donc de faire comprendre la théorie qui les décrit –, c'est parce que les modes de raisonnement ont évolué de concert avec le progrès des sciences et que les modèles mécaniques sont en conséquence ceux qui leur sont le plus adaptés ; rien ne permet d'affirmer qu'il en sera toujours ainsi.

le type de modèle considéré satisfaisant à un moment donné est fonction du climat intellectuel qui prévaut. Même si les modèles actuels de la théorie quantique nous semblent étranges et même « inintelligibles » ce n'est pas une raison contraignante pour affirmer que cette étrangeté ne va pas se dissiper avec une familiarité accrue, ou qu'une interprétation plus satisfaisante de la théorie ne sera pas finalement trouvée. De plus, l'inintelligibilité présumée du modèle présent provient en grande partie d'une inattention au fait que les mots « onde » et « particule » utilisés pour le décrire sont utilisés de manière analogique. (Nagel, 1961, p. 144)

Affirmer qu'une théorie n'a pas de signification tant qu'un modèle ne lui est pas fourni, et, en outre, que le modèle est une composante essentielle des théories, c'est adopter une conception de la notion de théorie qui se démarque de celle de Carnap au même titre que la conception des théories véhiculée par les physiciens mécanistes⁵⁸ se démarque de celle défendue par Duhem (1914). Pour Duhem comme pour Carnap (qui dispose des outils nécessaires à l'expression formelle de la conception duhémienne d'une théorie comme ensemble de symboles abstraits), *TC* est suffisante ; pour Nagel, comme pour les physiciens mécanistes, elle ne dit rien tant qu'elle n'est pas présentée dans un modèle. Ce point est de toute première importance pour la suite de mon propos : comme on le verra, admettre qu'une théorie, si on la restreint à *TC*, n'a pas de signification, c'est admettre les limites de la formalisation telle que l'envisagent les empiristes logiques.

Avant d'en venir à cette conclusion, une question se pose : est-il nécessaire, pour une caractérisation convenable de la notion de théorie, de faire des modèles une composante essentielle des théories ? On a vu que les modèles constituent une aide psychologique indispensable et occupent manifestement une place importante dans les raisonnements des scientifiques. Cela suffit-il pour autant à justifier qu'ils figurent dans une reconstruction formelle, au même titre que le squelette syntaxique et les règles de correspondance ? L'ambition initiale de Carnap, on l'a vu, consiste au contraire à se servir des outils formels pour se débarrasser de ces aspects psychologiques.⁵⁹

⁵⁸Je désigne par là les physiciens qui, comme Maxwell ou Kelvin, favorisent la présentation des théories au moyen de modèles mécaniques, et non des individus qui défendraient l'hypothèse atomiste (voir chapitre 4, page 172).

⁵⁹Dans un paragraphe intitulé « Understanding in Physics », Carnap (1939a, pp. 67-68) affirme clairement que les modèles n'ont pas leur place dans une reconstruction formelle : « Pendant longtemps, on n'avait pas conscience de la possibilité et même de la nécessité d'abandonner la recherche d'une compréhension [intuitive]. Quand des formules abstraites, non intuitives, comme, par exemple, les équations de l'électromagnétisme de Maxwell, étaient proposées en guise de nouveaux axiomes, les physiciens tentaient de les rendre "intuitives" en construisant un "modèle", c'est-à-dire, une manière de représenter les microprocessus électromagnétiques par analogie avec des processus macroscopiques connus, comme par exemple les mouvements des choses visibles. Plusieurs tentatives ont été

Insister sur l'importance de cette dimension de l'activité scientifique est tout à fait légitime⁶⁰, mais pourquoi vouloir l'intégrer à une reconstruction formelle ? La suite de cette section est destinée à donner des éléments de réponse à ces questions.

1.5.3 Le dédoublement de la sémantique

Avant d'évaluer l'importance et le statut des modèles en tant que composante essentielle des théories, revenons un instant sur les règles de correspondance et inversons la question : si ce sont les modèles qui permettent d'interpréter le squelette syntaxique de la théorie, si ce sont eux qui font que la théorie dit quelque chose, à quoi servent les règles de correspondance, que Carnap concevait précisément comme des règles d'interprétation ?

Il doit cependant être clair, malgré l'utilisation d'un modèle pour énoncer une théorie, que les hypothèses fondamentales de la théorie ne donnent que des hypothèses implicites à propos des notions théoriques qu'elles emploient. [...] De plus, la présentation d'une théorie par le biais d'un modèle ne rend pas moins impérieuse la nécessité de règles de correspondance pour relier la théorie à des concepts expérimentaux. Bien que les modèles des théories aient des fonctions importantes dans l'enquête scientifique, comme il sera montré dans le chapitre suivant, les modèles ne sont pas des substituts des règles de correspondance. [...] La distinction entre un modèle (ou interprétation) pour une théorie et les règles de correspondance pour les termes de la théorie est une distinction cruciale, et doit par conséquent être analysée plus avant (Nagel, 1961, p. 95).

faites dans cette direction, sans résultats satisfaisants. Il est important de prendre conscience du fait que la découverte d'un modèle n'a pas de valeur autre qu'esthétique ou didactique, ou encore, au mieux, heuristique, et qu'elle n'est pas du tout essentielle au succès de l'application de la théorie physique. L'exigence d'une compréhension intuitive des axiomes a été de moins en moins satisfaite quand les développements de la science ont conduit à la théorie de la relativité générale et à la mécanique quantique, impliquant la fonction d'onde. Beaucoup de gens, y compris des physiciens, ont un sentiment de regret et de déception à cet égard. Certains, en particulier des philosophes, vont jusqu'à prétendre que ces théories modernes, puisqu'elles ne sont pas intuitivement compréhensibles, ne sont pas du tout des théories à propos de la nature mais de "pures constructions formelles", "de purs calculs". Mais c'est une incompréhension fondamentale de la fonction d'une théorie physique. Il est vrai qu'une théorie ne doit pas être un "pur calcul" et doit posséder une interprétation, sur la base de laquelle elle peut être appliquée aux faits de la nature. Mais il est suffisant, comme on l'a vu, de rendre cette interprétation explicite pour les termes élémentaires ; l'interprétation des autres termes est donc indirectement déterminée par les formules du calcul, qu'il s'agisse de définitions ou de lois, les reliant aux termes élémentaires. » Je remercie Delphine Chapuis-Schmitz de m'avoir indiqué ce passage.

⁶⁰C'est une des préoccupations majeures de Kuhn, qui est justement compté comme un critique du positivisme logique et des approches formelles.

Le modèle de la théorie, s'il fournit une interprétation aux axiomes de la théorie et, ce faisant, les rend intelligibles – ou sensibles –, n'établit, selon Nagel, aucun lien entre eux et l'expérience. Plusieurs modèles peuvent être utilisés pour interpréter une théorie, faisant référence à des domaines de phénomènes différents, et pouvant entretenir certaines différences structurelles ; au mieux un seul d'entre eux est une représentation adéquate du monde. Le modèle d'une théorie, qui en expose le contenu, n'est pas présenté comme une représentation vraie des phénomènes. Illustrer les axiomes de la théorie cinétique des gaz à l'aide du modèle des boules de billard ne revient pas à affirmer que les molécules *sont* des boules de billard ; en donnant le modèle d'une théorie physique, on n'affirme pas plus qu'elle est vraie du monde qu'on ne l'affirme d'une théorie mathématique en en donnant un modèle visualisable. Ainsi, selon Nagel, si le modèle garantit *l'intelligibilité* de la théorie, les règles de correspondance en assurent le caractère *empirique*.

Nagel creuse donc un fossé entre la signification d'une théorie – ce qu'elle dit – et ses conséquences empiriques, alors que, on l'a vu, le projet initial de Carnap était, avec la théorie vérificationniste, de définir la signification cognitive d'un énoncé par ses conditions empiriques de vérité. Nagel, parce qu'il se situe dans la tradition empiriste logique qui consiste à reconstruire formellement une théorie en distinguant son squelette syntaxique des règles qui en permettent l'application empirique, et qu'il exige en même temps d'une théorie qu'elle dise quelque chose de plus qu'un simple instrument de prédiction, est conduit à opérer un *dédoublement de la sémantique* des théories : elles reçoivent leur interprétation *empirique* des règles de correspondance, mais cette interprétation ne leur donne pas réellement une signification. Pour Carnap, les énoncés d'une théorie ont une signification pour autant que l'on peut établir leurs conditions de vérité empirique ; leur « surplus de sens », celui qui n'est pas donné par l'interprétation partielle, n'entre pas dans la caractérisation de leur signification cognitive. Pour Nagel, la théorie fournit une interprétation sémantique à ses énoncés en décrivant des objets qui, si la théorie est vraie, sont une représentation adéquate du monde. C'est donc en quelque sorte par un dédoublement de l'interprétation de la théorie que Nagel comble le vide laissé par l'aspect incomplet de l'interprétation fournie par les règles de correspondance.

La tension entre la fonction des règles de correspondance et celle du modèle d'une théorie est cependant révélatrice du fait que la conception des théories qui sous-tend la reconstruction nagélienne n'est pas compatible avec la formalisation telle que les empiristes logiques la conçoivent. Dans la deuxième partie de ce chapitre, on verra que la conception sémantique est née, entre autres, de l'élimination de cette tension par le renoncement à l'idée selon laquelle on peut distinguer la partie syntaxique d'une théorie scientifique de son interprétation, et, corrélativement, à l'idée selon

laquelle le contenu d'une théorie est son contenu empirique. Auparavant, voyons les conséquences de cette position pour le traitement de ce que Nagel appelle le « statut cognitif des théories ». L'examen de ce problème, qui est celui de la légitimité de l'attribution d'une valeur de vérité aux énoncés d'une théorie et, corrélativement, de la nature et de l'existence des entités auxquels les concepts qu'elle emploie sont censés référer, va me permettre de mettre en évidence le caractère intenable du postulat fondamental de l'entreprise formaliste des empiristes logiques.

1.6 Le statut cognitif des théories

Dans ce qui précède, on a vu comment les empiristes logiques, en particulier Carnap (1956b, 1966) et Nagel (1961), proposent de reconstruire le contenu des théories. La théorie vérificationniste de la signification définissait la signification d'un énoncé par ses conditions empiriques de vérité : une théorie est donc, dans cette perspective, vraie ou fausse si et seulement si ses conséquences observationnelles le sont. Avec la thèse de l'interprétation partielle et l'admission d'un « surplus de sens » non empirique des termes théoriques se pose naturellement la question de la nature et de l'existence des entités désignées par les termes théoriques, ainsi que de la vérité de ce que la théorie dit à leur propos. En quel sens peut-on dire que les énoncés d'une théorie sont vrais ou faux ? Cette dernière question est celle de ce que Nagel appelle le « statut cognitif⁶¹ » des théories.

1.6.1 Instrumentalisme, réalisme et descriptivisme

Les trois conceptions concurrentes du statut cognitif des théories présentées par Nagel sont les conceptions instrumentale, descriptive et réaliste des théories. L'instrumentalisme, comme son nom l'indique, consiste à considérer que la théorie est un simple instrument destiné à tirer des prédictions empiriques, une boîte noire qui ne dit rien, n'affirme rien, et au sujet de laquelle il n'y a par conséquent aucun sens à demander si elle est vraie ou fausse.

Le réalisme, à l'opposé, consiste à soutenir qu'une théorie fait référence à des entités inobservables et que, par conséquent, s'interroger sur la vérité d'une théorie revient à s'interroger sur l'existence des entités qu'elle postule. Selon cette conception, le but des théories est de fournir la représentation la plus fidèle possible des entités

⁶¹ Attention, ici aussi, au sens de l'adjectif « cognitif » : tout comme dans l'expression carnapienne « signification cognitive », il ne désigne pas les processus mentaux des agents. Cependant, le « statut cognitif » ne doit pas être confondu avec la « signification cognitive ». La question du statut cognitif d'une théorie n'est en effet pas tant celle de sa signification, que de l'engagement – de l'attitude propositionnelle – dont elle est censée faire l'objet.

non observables. Elle n'est donc couronnée de succès – vraie – que si ces entités existent.

La conception descriptive, attribuée par Nagel, entre autres, à Pierre Duhem⁶², considère que les théories affirment des choses à propos des phénomènes observables, et peuvent par conséquent être vraies ou fausses. Cependant, tout ce qu'elles disent est contenu dans leurs conséquences observables ; toutes les autres choses dont elles ont l'air de parler ne sont qu'un effet de l'usage de symboles aveugles destinés à faciliter les prédictions. Elle est apparemment à mi-chemin entre les conceptions réaliste et instrumentale. Selon cette conception,

[...] une théorie est une formulation concise mais elliptique des relations de dépendance entre les événements et les propriétés observables. Bien que les affirmations d'une théorie ne puissent pas à proprement parler être dites vraies ou fausses quand elles sont prises au pied de la lettre, une théorie peut cependant être caractérisée ainsi dans la mesure où elle est traduisible dans des énoncés qui portent sur de l'observable. Les défenseurs de cette conception soutiennent, par conséquent, que dans le sens où une théorie (comme la théorie atomique) peut être dite vraie, les termes théoriques comme « atome » sont simplement des notations abrégées pour un complexe d'événements et d'aspects observables, et ne font référence à aucune réalité physique inaccessible à l'observation. (Nagel, 1961, p. 118)

1.6.2 Le descriptivisme bien compris est un instrumentalisme

La position descriptiviste ressemble fort à celle de l'empirisme logique orthodoxe, selon laquelle le contenu cognitif d'une théorie se réduit à son contenu empirique. Cependant, affirmer que les énoncés de la théorie parlent bien des phénomènes observables repose sur l'hypothèse de la *traductibilité* des énoncés théoriques en énoncés observationnels. Or, comme on l'a vu, le fait que les règles de correspondance ne fournissent qu'une interprétation partielle des termes théoriques interdit de les considérer comme des définitions explicites ou des règles de traduction. En conséquence, la position descriptiviste n'est tenable qu'à condition de considérer que l'interprétation fournie par les règles de correspondance est complète, et que la théorie n'est rien de plus qu'un simple compte rendu d'observation ; on a vu, à la section 1.2 du présent chapitre, qu'une telle position est insatisfaisante, et ne permet pas de rendre compte d'aspects essentiels de la notion de théorie, comme par exemple le fait que les théories permettent de découvrir de nouvelles lois empiriques. Si l'on refuse une

⁶²Nagel souligne le rôle que cette conception, associée à l'idée selon laquelle une théorie doit se contenter de décrire sans prétendre expliquer les phénomènes, a joué dans le développement de l'énergétisme contre la philosophie atomiste (Nagel, 1961, pp. 118-199 et note 9 p. 199).

telle conception réductrice de la notion de théorie et que l'on admet que les règles de correspondance ne fournissent qu'une interprétation partielle des énoncés théoriques, alors on doit, selon Nagel, renoncer à la position descriptiviste : comme je l'ai mentionné précédemment, les règles de correspondance, parce qu'elles ne sont pas des règles de traduction, ne transforment pas les énoncés du squelette syntaxique en énoncés doués de sens (Nagel, 1961, pp. 122-128). Elles permettent seulement de tirer des prédictions de cette « boîte noire ».

Par conséquent, Nagel affirme que la position descriptiviste – qui considère qu'une théorie ne dit rien d'autre que ce que l'on peut en tirer empiriquement – n'est pas tenable (à moins de concevoir une théorie comme un simple compte rendu d'observation) et se réduit à la position instrumentaliste, pour laquelle la théorie ne dit strictement rien. Puisqu'elle ne dit rien, cela n'a aucun sens de se demander si ce qu'elle dit est vrai.

Il s'ensuit que, pour la conception du statut cognitif des théories que l'on vient d'examiner, la vérité et la fausseté ne peuvent être attribuées à aucune théorie physique actuelle – du moins pas tant que sa traductibilité dans le langage observationnel est établie. Par conséquent, la conception que l'on étudie coïncide avec la seconde position mentionnée précédemment, selon laquelle les théories doivent être considérées comme des instruments pour mener l'enquête, plutôt que comme des énoncés à propos desquels les questions de vérité et de fausseté peuvent être soulevées de façon féconde. (Nagel, 1961, pp. 128-129)

Les énoncés des théories ainsi conçues ont la « *forme grammaticale* d'énoncés, mais la plupart de ces expressions ne sont pas, en fait, des énoncés, mais des *formes d'énoncés*. » (p. 132) La théorie « fonctionne comme un “principe directeur” ou un “ticket inférentiel” *en vertu duquel* on peut tirer des conclusions sur les faits observables, mais pas comme une prémisse à *partir de laquelle* ces conclusions sont obtenues » (pp. 129-130). Ainsi, pour Nagel, le fait que les énoncés non interprétés du squelette syntaxique ne se voient pas fournir une interprétation complète par les règles de correspondance empêche de les considérer comme les prémisses d'un raisonnement. La position descriptiviste implique donc de considérer la théorie tout entière comme une sorte de machine à calculer, dont aucun rouage ne signifie quoi que ce soit⁶³, mais qui permet, à partir d'un certain nombre d'énoncés doués de sens, de tirer un autre ensemble d'énoncés doués de sens (dans les deux cas, ce sont des énoncés d'observation). Il n'y a par conséquent aucune relation d'implication logique entre les énoncés de la théorie (qui ne sont pas à proprement parler des énoncés) et les énoncés d'observation.

⁶³Pas plus, pour reprendre l'exemple de Nagel, qu'un marteau ne « représente » les choses que l'on a construites par son moyen (p. 130).

La théorie moléculaire des gaz n'est donc pas plus impliquée logiquement (selon les défenseurs de la conception instrumentaliste)⁶⁴ qu'elle n'implique des énoncés à propos de faits d'observation. La *raison d'être*⁶⁵ de la théorie est de servir de règle ou de guide pour faire des transitions logiques d'un ensemble de données expérimentales à un autre. (Nagel, 1961, pp. 129-130)

Reprenons la conclusion à laquelle arrive Nagel. Elle repose sur l'hypothèse *H* selon laquelle les règles de correspondance ne sont pas des règles de traduction (si l'on refuse cette hypothèse, on renonce à faire d'une théorie autre chose qu'un compte rendu d'observation). De *H*, Nagel arrive à la conclusion *C* que le descriptivisme se réduit à l'instrumentalisme. En effet, le descriptivisme, selon lequel la signification d'une théorie est sa signification empirique, est intenable, car il repose sur une hypothèse fautive (celle selon laquelle les règles de correspondance sont des règles de traduction). Dès lors que l'on accepte *H*, on doit reconnaître *C*, et donc admettre que la théorie n'a pas de signification. Cependant, Nagel *n'affirme pas que l'instrumentalisme s'impose*, puisque, précisément, *H*, en vertu de quoi il montre *C*, autorise à considérer que *le contenu de la théorie ne se réduit pas à son contenu empirique*. Le « surplus de sens » laissé de côté par l'interprétation empirique partielle, s'il n'entre pas, pour Carnap, dans la caractérisation du contenu cognitif de la théorie, est au contraire de toute première importance aux yeux de Nagel : c'est dans ce « surplus de sens » que réside toute la signification de la théorie, puisque, précisément, son interprétation empirique ne lui fournit aucune signification – en vertu de *H*.

Ainsi, Nagel établit que, *en ce qui concerne son contenu empirique*, (*RTC* pour Carnap, les prédictions permises par le squelette syntaxique et les règles de correspondance pour Nagel), la théorie ne dit strictement rien et que cela n'a aucun sens de demander si elle est vraie ou fautive : comme un instrument, elle fonctionne ou ne fonctionne pas. Si, comme Carnap, on considère que le contenu cognitif d'une théorie se réduit à son contenu empirique, alors on doit adopter une position instrumentaliste selon laquelle la théorie ne dit rien. Mais on peut refuser cette hypothèse, et considérer que la théorie a un contenu de signification non empirique. Dans ce cas, rien ne nous contraint à adopter une position instrumentaliste.

1.6.3 Les modèles ne disent rien du monde

Revenons un instant à la conception de Carnap : s'il affirme que le contenu cognitif des théories se réduit à leur contenu empirique (leur énoncé de Ramsey *RTC*), il

⁶⁴Et donc, selon les défenseurs de la position descriptiviste, puisque Nagel montre que cette dernière se réduit à la position instrumentaliste.

⁶⁵En français dans le texte.

reconnaît que la théorie ne se réduit pas à son contenu empirique, au sens où l'énoncé de Ramsey de la théorie n'est pas la théorie (voir page 208). C'est essentiellement par la manière dont ils rendent compte du surplus de sens laissé de côté par l'interprétation partielle que Carnap et Nagel divergent. Comme on va le voir, ce que Carnap considère comme un effet du langage que l'on s'est donné pour exprimer le contenu cognitif de la théorie constitue, pour Nagel, la signification même de la théorie, qui lui est fournie par son modèle.

La question qui se pose ici est celle du statut cognitif de ce « surplus de sens ». La reconnaissance du fait que l'on ne peut pas réduire une théorie à la mise en forme symbolique d'un compte rendu d'observation risque-t-elle de conduire le philosophe à renoncer à l'empirisme et à admettre l'existence de mystérieuses entités ? Moins dramatiquement, est-on obligé, une fois ce fait reconnu, d'adopter une position réaliste et d'affirmer que, dans la mesure où la théorie atteint son but, les entités dont elle parle existent ?

La solution de Ramsey-Carnap, au contraire, est destinée à montrer que la partie non empirique de la théorie, son postulat de signification A_T , *n'est pas une assertion*. Si la théorie de l'électricité *semble dire* quelque chose de plus à propos des électrons que ce qui est contenu dans son énoncé de Ramsey, c'est un simple effet du langage que l'on utilise pour la formuler. Ce langage nous fait indûment croire que la théorie parle d'entités réelles que l'on appelle « électrons », dont on n'a pas une définition complète, et à propos desquelles le but de la science est d'établir de nouvelles vérités. En fait, A_T (la partie non empirique de la théorie) est un énoncé analytique qui confère un sens au terme « électron ». La présentation de la théorie sous la forme de Ramsey-Carnap dissipe ainsi, selon Carnap, l'illusion selon laquelle la théorie dit quelque chose de plus à propos des électrons que ce qui est contenu dans son énoncé de Ramsey :

Dans la manière dont Ramsey propose de parler du monde extérieur, un terme comme « électron » disparaît. Cela n'implique aucunement que les électrons disparaissent, ou, plus précisément, que tout ce qui, dans le monde extérieur, est désigné par le mot « électron » disparaisse. L'énoncé de Ramsey continue d'affirmer, au moyen de ses quantificateurs existentiels, qu'il y a quelque chose dans le monde extérieur qui possède toutes les propriétés que les physiciens attribuent à l'électron. Il ne remet pas en question l'existence – la « réalité » – de ce quelque chose. Il propose simplement une autre manière d'en parler. La question difficile qu'il évite n'est pas : « les électrons existent-ils ? », mais « quelle est l'exacte *signification* du terme "électron" ? » Dans la manière dont Ramsey propose de parler du monde, cette question ne se pose pas. Il n'est plus nécessaire d'enquêter sur la signification d'« électron », puisque le terme lui-même n'apparaît pas dans le langage de Ramsey. (Carnap, 1966, chap. 26,

p. 252)

En somme, la question « Les électrons existent-ils ? » n'est qu'une question linguistique déguisée. Ou bien c'est une question interne au cadre linguistique dans lequel on se situe, et dans ce cas la réponse est triviale : lorsque je me situe dans le cadre de la théorie de l'électricité et que je cherche à décrire le comportement d'un certain courant électrique, la réponse est évidemment positive.⁶⁶ Ou bien c'est une question externe au cadre linguistique. Dans ce cas, s'il s'agit d'une question métaphysique sur l'existence dans le monde d'un certain type d'entités que l'on appelle « électrons » et qui jouent le rôle qu'on assigne au terme théorique « électron » dans la théorie de l'électricité, cette question n'a pas de signification cognitive.

Cette dernière sorte de question peut la plupart du temps être correctement reformulée sous la forme d'une question pragmatique sur le choix du langage le plus commode : est-il plus commode d'adopter un langage dans lequel on peut quantifier sur une classe d'objets appelés « électrons » ? Or, en matière de choix de langage, le conventionnalisme de Carnap nous autorise à toute la prodigalité jugée utile⁶⁷ : on compliquera au pire inutilement les règles de calcul, mais sans danger pour l'ontologie impliquée par la théorie (Carnap, 1956b, p. 42).⁶⁸ Pourvu que l'inflation de l'appareil logique soit strictement explicitée à l'aide de quantificateurs existentiels, elle n'entraîne selon Carnap aucune inflation ontologique (les variables liées aux quantificateurs ne sont pas du même type logique que les anciennes constantes qui représentaient les termes théoriques) ; le langage logique que l'on se donne n'implique pas d'autre existence que celle des référents des constantes – par conséquent des observables.⁶⁹ Ainsi, une question comme « Les électrons existent-ils », pour autant qu'elle

⁶⁶Cela ne revient pas du tout à dire que toutes les questions relatives aux électrons dans le cadre d'une théorie sont triviales : déterminer la charge d'un électron particulier ou le mouvement des électrons dans tel atome, par exemple, ne sont pas des tâches triviales, et constituent précisément l'objet du travail scientifique.

⁶⁷Le « principe de tolérance » défendu par Carnap consiste en effet à affirmer qu'« *en logique, il n'y a pas de morale*. Chacun est libre de construire sa propre logique, i.e. sa propre forme de langage, comme il le désire. Tout ce que l'on exige de lui est que, s'il veut qu'on en discute, il indique clairement la manière dont il entend procéder, et fournisse des règles syntaxiques plutôt que des arguments philosophiques. » (Carnap, 1934b, §17)

⁶⁸C'est là un des points de désaccord fondamentaux entre Carnap et Quine. Pour Quine (1953b), l'adoption d'un certain langage implique un engagement ontologique quant à l'existence des classes d'entités désignées par les variables liées (les variables sur lesquelles le langage autorise à quantifier). Pour Carnap (1956a), en revanche, l'adoption d'un langage est une affaire pragmatique, qui n'implique aucun engagement ontologique quant à ces classes d'entités.

⁶⁹Soulignons par ailleurs que Carnap admet explicitement que l'appareil non descriptif de ce langage peut inclure toutes les mathématiques nécessaires à l'expression des théories de la science empirique : « D [c'est-à-dire le domaine d'entités admises comme valeurs des variables de L_T] inclut

a une signification cognitive, peut être reposée sous la forme « Doit-on préférer un langage qui emploie le terme “électron” ? »

On retrouve ici une conviction carnapienne qui remonte à ses tout premiers travaux : toutes les questions métaphysiques – ou toutes les questions qui semblent porter sur le monde mais que la science ne peut pourtant pas résoudre car elles ne peuvent pas recevoir de réponse empirique – sont des questions linguistiques mal posées.⁷⁰ Une fois correctement posées, elles se réduisent à des questions de choix pragmatiques, que le logicien philosophe ne peut pas trancher au moyen d'outils formels : leur réponse dépend des buts poursuivis par les scientifiques.

De cela, Carnap (1966) conclut donc que *le débat entre instrumentalisme et réalisme n'a pas à être tranché de manière formelle et qu'il se réduit à une question de choix de langage*. Comme je le montrerai à la section suivante, si le débat entre instrumentalisme et réalisme est bien une affaire pragmatique, ce n'est pas en tant qu'il se réduit à la question du *choix* d'un langage qu'à celle de la manière dont on *utilise* et dont on *comprend* le langage qu'on s'est donné. Auparavant, voyons comment Nagel (1961), auquel Carnap (1966, p. 256, note 3) fait explicitement référence sur ce point, soulignant la concordance de leurs vues, en arrive lui-même à la conclusion selon laquelle le débat entre instrumentalisme et réalisme est une affaire pragmatique.

Le divorce prononcé entre les conséquences empiriques de la théorie et son contenu

tous les types d'entités dont on a besoin pour la partie purement mathématique de L_T . » (Carnap, 1956b, p. 43) Carnap n'affirme pas ici que les théories mathématiques elles-mêmes doivent être formalisées au premier ordre. Il semble donc que la caractérisation que les défenseurs de la conception sémantique proposent du projet de formalisation des empiristes logiques soit historiquement inexacte (comme je l'ai signalé précédemment, note 7). Je reviendrai sur ce problème à la section 2, page 245.

⁷⁰ « Il n'y a rien de plus, rien de “plus élevé” à dire au sujet des choses que ce que la science en dit. L'objet de la logique de la science est plutôt la science elle-même en tant que structure ordonnée de propositions. » (Carnap, 1934a, p. 196) « Ou bien le travail scientifique a pour objet le contenu empirique des propositions : on observe, expérimente, collecte et élabore le matériau de l'expérience; ou bien il s'agit de mettre au clair la forme des propositions de la science » (Carnap, 1932a, p. 322). Avant d'admettre la sémantique dans son système (voir ci-dessus, note 21), Carnap proposait une distinction présentée comme syntaxique entre idiome formel et idiome matériel (ou contentuel) : les questions qui ont l'air de porter sur le monde (parce qu'elles sont exprimées au moyen de l'idiome matériel) mais auxquelles la science ne peut pas répondre sont en fait des questions linguistiques déguisées. Elles doivent alors pouvoir être reformulées en des questions sur le langage (employant l'idiome formel). Voir (Carnap, 1934a) et (Carnap, 1934b, § 73 à 81). Ainsi, par exemple, la proposition du discours matériel « les *nombres* sont des classes de classes de choses », correctement reformulée au moyen de l'idiome formel, donne : « Les expressions numériques sont des expressions de classe du second ordre. » (Carnap, 1934b, §78). De même, la proposition « le *temps* est uni-dimensionnel; l'*espace* est tri-dimensionnel. » peut être adéquatement reformulée comme suit : « Une désignation de temps est constituée d'une seule coordonnée; une désignation spatiale est constituée de trois coordonnées. » (§ 79).

sémantique (voir ci-dessus, section 1.5.3) lui a permis de dire que si la présentation, l'adoption et l'utilisation d'une théorie sont indissociables d'un modèle qui lui donne de la « chair », ce modèle n'est pas pour autant censé être une représentation vraie du monde (voir ci-dessus, page 220). Par conséquent, adopter la théorie cinétique des gaz n'engage aucunement à croire que les molécules de gaz *sont* des boules de billard.

La seule chose que l'on puisse affirmer avec confiance, c'est qu'un modèle d'une théorie n'est pas la théorie elle-même. En conséquence, on ne peut pas considérer sans un examen plus poussé que l'adéquation de la théorie comme instrument pour l'explication et la prédiction systématiques établisse la réalité physique de chaque aspect du modèle matériel dans les termes duquel la théorie peut être interprétée. Cela est évident dans le cas où il existe plusieurs modèles connus pour la même théorie, mais cela est également vrai quand on ne dispose que d'un seul modèle (Nagel, 1961, p. 116).

Les modèles sont indispensables à la construction, au développement et à la formulation des théories; cependant, comme on l'a vu, le type de modèle qu'on utilise dépend de facteurs contextuels et psychologiques (voir page 217). Ces facteurs peuvent conduire à construire des modèles qui contiennent de nombreux éléments superflus, et parfois trompeurs. L'utilisation de modèles imagés et suggestifs peut présenter un danger pour le théoricien qui risque d'être entraîné sur de fausses pistes suggérées par des éléments inessentiels du modèle qui n'ont aucun référent dans le monde réel. Comme le remarquait Hertz, les images mentales que nous fournissent nos théories scientifiques portent la marque des esprits qui les produisent, et sont par conséquent susceptibles de contenir des éléments superflus :

[Hertz] reconnaissait clairement que cette exigence instrumentale ne détermine pas de manière unique le symbolisme (ou la théorie) qui remplira cet objectif. Il notait, en particulier, qu'une théorie contiendra inévitablement ce qu'il appelait des « relations superflues ou vides » – des symboles qui ne représentent rien dans le domaine de phénomènes pour lequel la théorie est construite. Selon Hertz, ces « relations vides » entrent dans nos théories simplement parce que les théories sont des symboles complexes, « des images produites par notre esprit et nécessairement affectées par les caractéristiques de sa manière de représenter. » (Nagel, 1961, p. 103).

Il est indispensable, pour construire et utiliser une théorie, qu'elle soit présentée dans un certain modèle, sans quoi, on l'a vu, elle est inintelligible, mais cela n'engage aucunement, en principe, son utilisateur, à affirmer que ce modèle est vrai (qu'il représente adéquatement le monde). Par conséquent, l'argumentation de Nagel – telle que je la reconstruis – le conduit à établir les deux thèses suivantes :

1. En ce qui concerne son contenu empirique, la théorie ne dit rien (voir ci-dessus, page 224), et ne peut donc pas être dite vraie ou fausse.

2. Ce que l'on appelle son « surplus de sens », n'est aucunement un contenu factuel ; il n'est donc pas plus susceptible d'être vrai ou faux (il est vrai du modèle, mais cela ne signifie pas qu'il soit vrai du monde).

1.6.4 Les théories scientifiques sont toujours déjà interprétées

Comme je l'ai dit, ces considérations conduisent Carnap, qui se réclame sur ce point de Nagel, à conclure que le débat entre réalisme et instrumentalisme est « essentiellement d'ordre linguistique. Il s'agit de savoir quelle manière de parler est préférable dans un contexte donné » (Carnap, 1966, chap. 26, p. 256). Que signifie précisément cette affirmation ? Que veut dire, ici, « manière de parler » ? Pour le comprendre, soulignons d'abord que Carnap ne prétend naturellement pas que les théories gagneraient à être présentées dans la formalisation de Ramsey, à supposer qu'elles le puissent : les scientifiques ont raison d'utiliser des formulations qui emploient des termes comme « électrons », parce qu'elles sont beaucoup plus commodes. Il s'agit cependant, selon lui, de rester conscient de la distinction entre le contenu empirique de la théorie et les effets du langage utilisé, choisi pour sa commodité.

Les physiciens, certes, trouvent beaucoup plus commode de parler dans un langage abrégé qui inclut des termes théoriques comme « proton », « électron » et « neutron ». Mais si on leur demande si les électrons « existent » réellement, ils répondront certainement de diverses façons. Certains physiciens trouveront satisfaisant de concevoir des termes comme « électron » suivant la manière de Ramsey. Ils esquivent la question de l'existence en affirmant qu'il y a des événements observables, par exemple dans des chambres à bulles, qui peuvent être décrits au moyen de certaines fonctions mathématiques, dans le cadre d'un certain système théorique. Ils n'affirmeront rien de plus. Demander s'il *existe* réellement des électrons revient au même – du point de vue de Ramsey – que demander si la physique quantique est vraie. La réponse est que, dans la mesure où la physique quantique a été confirmée par des tests, on est autorisé à dire qu'il existe des cas où se produisent certains types d'événements qui, dans le langage de la théorie, sont appelés « électrons ». (Carnap, 1966, chap. 26, pp. 54-55)

Cependant, cette distinction, faite par le physicien très philosophe décrit ci-dessus, entre les conséquences empiriques de la théorie et ce que sa formulation dans un langage qui quantifie sur des variables symboliquement notées « électrons » suggère à notre imagination fautive, n'est généralement pas faite en pratique. Le scientifique employant un langage qui quantifie sur le terme « électron », quand bien même il serait disposé, comme le suggèrent les propos de Carnap, à concevoir ce terme suivant la méthode de Ramsey, l'emploie effectivement, en pratique, dans des énoncés qui jouent le rôle de prémisses dans ses raisonnements. On a vu précédemment que la position

instrumentaliste consistait à considérer les énoncés des théories comme des règles d'inférence plutôt que comme des prémisses ; Carnap et Nagel s'accordent à dire que l'on est toujours libre de considérer une règle d'inférence comme la prémisse d'un raisonnement, selon le contexte :

Ceux qui caractérisent les théories comme des principes directeurs, comme des règles suivant lesquelles les inférences sont faites plutôt que comme des prémisses à partir desquelles on peut dériver des conclusions, négligent souvent la nature contextuelle de cette distinction. (Nagel, 1961, p. 138)⁷¹

Autrement dit, la question de savoir si je dois considérer un énoncé qui contient le terme *électron* comme une règle d'inférence ou comme une prémisse (et donc une authentique assertion) est elle-même une question pragmatique. Cette question pragmatique n'est pas la même que celle, précédemment mentionnée, concernant le choix du langage le plus commode. Ici, *la question n'est pas celle du choix d'un langage mais celle de la manière dont on l'utilise*. La première question (celle du choix du langage) est la suivante : « Dois-je utiliser un langage qui quantifie sur le terme “électron” ? ». La seconde question (celle de l'utilisation du langage) est la suivante : « Dois-je considérer qu'une expression comportant le terme “électron” est une assertion ou une règle d'inférence ? »

C'est parce que cette seconde question est elle-même une question pragmatique que Carnap, comme Nagel, peuvent dire que l'instrumentalisme et le réalisme ne sont pas deux thèses philosophiques, mais correspondent à deux « manières de parler ». Ces deux « manières de parler » ne sont pas deux *langages* différents mais deux *attitudes épistémiques* différentes des utilisateurs de la théorie. Ces deux attitudes épistémiques, répétons-le, ne concernent pas le choix d'un langage, mais la manière dont on utilise et on comprend le langage de la théorie (quel qu'il soit) : soit on considère que les énoncés de la théorie fournissent des règles qui, par la manipulation d'un certain système symbolique, permettent de tirer des prédictions empiriques (auquel cas on a une attitude instrumentaliste), soit on utilise ces énoncés comme des prémisses dans nos raisonnements. Dans ce dernier cas, on a une attitude réaliste. En effet, si l'on considère les énoncés des théories comme les prémisses d'un raisonnement et non comme des règles d'inférence, on doit admettre qu'ils ont une signification, qu'ils disent quelque chose ; on est donc assez naturellement à se demander ce dont ils parlent et, par conséquent, ce que sont les entités dont ils parlent (par exemple, les électrons). De là à se demander si elles existent vraiment, il n'y a qu'un pas.

Or, remarque Nagel, de ce point de vue, la pratique semble généralement trancher en faveur d'une position réaliste :

⁷¹Voir aussi (Carnap, 1956b, p. 47).

De fait, les théories sont habituellement présentées et utilisées comme des prémisses, plutôt que comme des principes directeurs, dans les traités scientifiques ainsi que dans les articles rapportant les résultats de la recherche théorique et expérimentale. Certains parmi les scientifiques les plus éminents, vivants ou morts, ont très certainement considéré les théories comme des énoncés portant sur la constitution et la structure d'un domaine de phénomènes donné ; et ils ont conduit leurs recherches sur l'hypothèse que la théorie était une *carte projetée* d'un domaine de la nature, plutôt qu'un ensemble de *règles de projection*. Une grande partie de la recherche expérimentale est indubitablement inspirée par le désir de vérifier si les entités et processus hypothétiques postulés par la théorie [...] se produisent effectivement (Nagel, 1961, p. 139).

Notons bien ce que Nagel veut dire par là : non pas que le réalisme est la meilleure manière de rendre compte du statut cognitif des théories, mais que l'activité théorique elle-même se caractérise par une attitude épistémique globalement réaliste de la part des scientifiques. La construction d'hypothèses, leur développement et leur mise en forme mathématique, on l'a vu, nécessite que l'on raisonne sur des modèles dont le philosophe a bien montré qu'ils n'étaient pas affirmés du monde, mais dont le scientifique fait *comme s'ils l'étaient* pour parvenir à élaborer et à utiliser sa théorie.⁷²

Cette remarque se résume-t-elle à une considération historique sur la psychologie des chercheurs selon laquelle le réalisme est l'attitude la plus fructueuse du point de vue de la découverte et de l'invention théoriques ? On retrouve ici une question posée précédemment (voir page 219) : quelle pertinence ces remarques concernant le travail de découverte scientifique ont-elles pour la reconstruction formelle du contenu des théories déjà construites ? Si les modèles sont simplement une manière de rendre compte de la réalité psychologique à laquelle correspond le « surplus de sens » des termes théoriques – la partie non empirique de leur signification – quelle légitimité ont-ils pour figurer comme composante essentielle ?

C'est ici que réside à mon avis la clé de la position de Nagel et de son refus de voir dans les règles de correspondance une interprétation – même partielle – du squelette logique. C'est aussi, on va le voir, ce qui marque les limites de la reconstruction formelle des théories à la manière des empiristes logiques.

Dire que les théories sont habituellement prises comme des prémisses et que leur modèle en est une composante essentielle, c'est en fait dire que les théories sont *toujours déjà interprétées*. On l'a vu précédemment, la construction, le développement de l'appareil mathématique d'une théorie et son application à de nouveaux types de phénomènes ne peut pas se faire sans qu'elle soit présentée dans un certain modèle.

⁷²Sur cette question, voir aussi (Vaihinger, 1911) et (Fine, 1993).

L'utilisation d'analogies matérielles, comme celle des boules de billard, mais également le fait que la plupart des termes nouveaux introduits dans les théories sont empruntés au langage ordinaire ou forgés à partir d'une étymologie suggestive (« force », « électron », « quantum »), et enfin le simple fait d'utiliser un certain type d'équation, un certain langage mathématique, tout cela revient déjà à se situer dans un certain modèle (je veux dire par là que le modèle n'est pas une simple aide psychologique quand on n'arrive pas à comprendre une théorie : une théorie *comprise* est interprétée) :

La théorie de l'atome de Bohr n'est généralement pas présentée comme un ensemble abstrait de postulats, avec un nombre approprié de règles de correspondance pour les termes non logiques non interprétés implicitement définis par les postulats. Elle est habituellement exposée au moyen de notions relativement familières, de telle sorte que, au lieu d'apparaître comme des formes d'énoncés, les postulats de la théorie apparaissent comme des énoncés, dont le contenu peut être au moins partiellement visualisé. Une telle présentation est adoptée, entre autres raisons, parce qu'elle peut être comprise avec plus de facilité que ne le peut une exposition formelle inévitablement plus longue et compliquée. Mais, de toute manière, ainsi exposés, les postulats de la théorie sont ancrés dans un modèle ou interprétation. (Nagel, 1961, p. 95)

Ainsi, Nagel n'affirme pas que le physicien est un être naïf croyant sincèrement que les molécules du gaz sont adéquatement représentées par des boules de billard, mais que le simple fait de développer et d'utiliser une théorie requiert qu'elle soit présentée sous une certaine forme et que cette forme, ce langage que l'on utilise, nous situe d'emblée dans une interprétation. Par conséquent, les théories sont toujours présentées dans un modèle (et non optionnellement accompagnées d'un modèle).⁷³

C'est la raison pour laquelle Nagel parle généralement plutôt *du* modèle dans lequel une théorie est présentée que *des* modèles que l'on peut en donner, bien qu'une même théorie, si on la restreint à sa première composante, à savoir en tant qu'ensemble d'axiomes non interprétés, soit susceptible d'avoir une multiplicité – une infinité, en droit – de modèles.

On pourrait objecter que toutes ces remarques relèvent, encore et toujours, de considérations sur la découverte scientifique, et que l'impossibilité en pratique d'utiliser une théorie qui ne soit pas interprétée ne prouve pas que le philosophe a tort d'en proposer une reconstruction purement syntaxique et de vouloir l'identifier à sa structure syntaxique. En d'autres termes, le fait que les physiciens mécanistes aient

⁷³ « La plupart des théories, de toute façon, sont générées dans la matrice d'un certain modèle et sont codifiées, avec au mieux une mention en passant de règles de correspondance, dans les termes d'une interprétation de leurs prémisses fondamentales ». (Nagel, 1961, p. 107)

besoin de se représenter la théorie au moyen d'un modèle ne prouve pas que Duhem a tort de définir une théorie comme un ensemble de symboles abstraits (voir ci-dessus, page 218). La question qui se pose est la suivante : quand bien même il serait possible, au moins en principe, de présenter la mécanique classique, par exemple, sous la forme d'un squelette syntaxique et de règles de correspondance (*TC*), le résultat pourrait-il être appelé « théorie » ?

En premier lieu, l'analyse que je viens de proposer montre que si l'on tient, comme cela semble être le cas de Nagel mais aussi de Carnap (voir page 208), à faire d'une théorie un outil de prédiction, la réponse semble être « non » : il est difficile de comprendre en quoi consiste le « pouvoir prédictif » d'une théorie si celle-ci ne peut pas être utilisée pour tirer des prédictions. En outre, à supposer même que l'on parvienne à « tirer » des prédictions de cette boîte noire au moyen de recettes, une telle « théorie » serait strictement dépourvue de signification.⁷⁴ Cela signifie qu'elle serait non seulement inutilisable en pratique mais qu'elle serait dépourvue de ce qui fait d'elle une théorie proprement dite et non le simple enregistrement de rapports d'observation. L'artifice logique de Carnap qui permet de décrire le travail des scientifiques comme la manipulation de symboles non interprétés suivant certaines règles d'inférence permettant de générer des formules ayant une signification observationnelle ne résiste pas au constat suivant : l'utilisation d'une théorie ne consiste pas et ne peut pas consister en une telle manipulation qui se réduit à un calcul, à l'exécution d'un programme, au sens informatique. Pour reprendre les termes de Nagel, les scientifiques prennent les théories comme des *cartes projetées* plutôt que comme des *règles de projection* (Nagel, 1961, p. 139). Quand ils n'ont aucun moyen de se figurer une telle carte et en sont réduits à manipuler des formules sans les comprendre, il semble difficile d'affirmer que ces formules constituent une théorie.

Insistons : cette dernière remarque ne consiste pas seulement à dire que les théories sont, de fait, toujours présentées sous un certain habillage linguistique et mathématique et que la reconstruction formelle de son squelette syntaxique ne permet pas de rendre compte de la *pratique* scientifique. Cela est évident, et ne constitue pas, tel quel, une critique du projet des empiristes logiques, qui ne cherche pas à rendre compte des aspects pratiques de l'activité théorique. La distinction « contextuelle » entre prémisses et règles d'inférences recoupe en fait la distinction entre la reconstruction formelle de la théorie par le logicien et son utilisation par le scientifique ; or ce que la lecture de Nagel nous conduit à conclure, c'est que la formalisation, au sens où elle vise à dissocier la partie non interprétée de la théorie de son contenu,

⁷⁴Sauf, encore une fois, à considérer la théorie comme un simple compte rendu d'observation – ce qui est peut-être la position de Duhem, mais que les empiristes logiques, y compris Carnap, rejettent (voir section 1.2).

ne permet pas de retrouver ce contenu. Dire qu'une théorie, même formulée dans un langage hautement « abstrait » pour le profane et n'utilisant quasiment aucun terme du langage ordinaire, est toujours déjà interprétée, c'est dire que *la différence entre la formalisation du logicien et la formulation du scientifique n'est pas seulement une différence dans les outils linguistiques autorisés mais dans la manière dont on utilise ces outils.*

Le dilemme du théoricien (appelé aussi paradoxe de la théorisation), dont j'ai montré (section 1.2) qu'il exprimait le problème majeur que la reconstruction formelle des théories empiriques doit affronter, nous fait finalement arriver à la conclusion selon laquelle il y a quelque chose d'essentiel à la caractérisation de la notion de théorie qui résiste à la formalisation exigée par la recherche carnapienne d'un critère de signification cognitive. Une telle formalisation, qui consiste à distinguer entre le cœur axiomatique non interprété de la théorie et sa sémantique, nous fait inéluctablement manquer la partie du contenu de la théorie qui n'est pas son contenu purement factuel (empirique) et qui est pourtant le propre d'une théorie scientifique.

Vouloir enrégimenter le langage au point d'en éliminer tout terme qui pourrait suggérer des représentations mentales « superflues » en raison de sa signification extra-théorique c'est oublier que, même très ésotérique et éloigné du langage ordinaire, le langage d'une théorie est toujours utilisé en tant que langage interprété, armé d'une sémantique, au même titre que le langage naturel l'est.⁷⁵ Autrement dit, il est une partie du langage naturel, que seuls certains experts maîtrisent, mais qui ne jouit pas d'un statut sémantique d'un autre type.⁷⁶

Cette analyse autorise à conclure que les trois composantes dégagées par Nagel n'ont pas le même statut : apparemment distinguées par l'analyse des théories, elles n'en sont pas, en fait, des composantes au même titre. Les deux premières correspondent à la reconstruction rêvée du logicien, qui n'a pas de signification ; la troisième à la théorie interprétée. La tension que j'ai signalée à la section 1.5.3 entre modèle et règles de correspondance (ce que j'ai appelé le « dédoublement de la sémantique ») est le résultat, chez Nagel, de son attachement au projet empiriste de formalisation syntaxique des théories. Il contribue pourtant à marquer les limites d'un tel projet, en mettant en évidence le fait qu'une théorie qu'on dissèque en séparant sa syntaxe de sa sémantique n'est tout simplement plus une théorie scientifique.

⁷⁵Voir aussi, sur ce point, la critique de la notion de règle de correspondance par Schaffner (1969a).

⁷⁶Voir (Hempel, 1970) à propos de l'importance de l'utilisation, dans la formulation d'une théorie, de termes du langage naturel qui sont porteurs d'une signification extra-théorique, mais dont la signification est modifiée par leur utilisation dans le cadre de la théorie. Ces considérations s'accompagnent, chez Hempel, d'un rejet de la distinction entre langage observationnel et langage théorique, et, corrélativement, d'un grand scepticisme à l'égard de l'entreprise axiomatique.

1.6.5 Les formulations de la mécanique comme autant de modèles différents ?

Notons, avant de conclure, que la manière dont Nagel rend compte de la diversité des formulations de la mécanique est symptomatique de la tension que je viens de souligner dans sa pensée, elle-même significative de problèmes internes à la conception des théories défendue par les empiristes logiques. De prime abord, son analyse de la notion de modèle, et en particulier ses réflexions sur les modèles formels, qui autorisent à considérer que l'expression d'une théorie dans un certain langage mathématique porte, en elle-même, un modèle, semblent suggérer que *les différentes formulations de la mécanique en sont des modèles différents*.

En effet, comme on l'a vu page 232, un même squelette syntaxique peut se voir « habiller » par différents modèles, qui n'en modifient pas les conséquences empiriques, données par les règles de correspondance, mais en permettent une compréhension différente. Ainsi, une équation vectorielle de la forme du principe fondamental de la dynamique porte en elle un schéma d'explication mécaniste. Un principe variationnel en porte un autre. Les équations newtoniennes parlent de forces, les équations lagrangiennes et hamiltoniennes parlent d'énergie. En raison de leur expression dans un langage particulier, elles suggèrent une interprétation différente du squelette syntaxique qui, lui, à supposer qu'on puisse l'énoncer, ne parle ni de forces ni de d'énergie, mais consiste en un ensemble de symboles non interprétés directement reliés, par un ensemble de « recettes », au monde empirique.

Si l'on suit jusqu'au bout les propos de Nagel à propos du rôle des modèles dans l'explication (voir section 1.5.2), on doit donc admettre que les explications proposées par la mécanique sous ses différentes formulations ne sont pas les mêmes. Autrement dit, on est conduit à remettre en cause l'univocité de la notion de « modèle mécanique ». ⁷⁷ C'est précisément pour clarifier cette notion d'explication mécaniste que Nagel, au chapitre 7 de *The Structure of Science*, intitulé « Mechanical Explanations and the Science of Mechanics », entreprend de reconstruire le contenu de la mécanique classique. Or, comme on va le voir, Nagel écarte dès l'ouverture du chapitre le problème que je viens de poser, et qui semble naturellement découler de ses propos sur la notion de modèle.

La méthode la plus directe et satisfaisante d'établir le domaine d'une science et le caractère distinctif de ses explications est de se tourner vers les lois et théories générales – quand ces théories existent – qui constituent à une certaine étape de son développement les prémisses ultimes de ses explications. Heureusement, il est possible de faire cela dans le cas de la mécanique classique, car le contenu

⁷⁷ Voir ci-dessus, chapitre 4, note 16.

de cette science peut être assez bien saisi dans le cadre des idées que nous fournissent les « axiomes » ou « lois » newtoniens du mouvement. (Nagel, 1961, pp. 157-158)

Le projet de Nagel, ici, n'est pas de dégager le squelette syntaxique de la théorie. Bien au contraire, son entreprise de clarification conceptuelle de la mécanique classique repose sur le choix d'une de ses formulations linguistiques, la formulation newtonienne.⁷⁸ Cependant, Nagel semble considérer que les conclusions auxquelles l'axiomatisation (et non la formalisation au sens strict, puisque cette axiomatisation repose sur l'adoption d'un certain langage mathématique) de la mécanique *newtonienne* vont le conduire peuvent être généralisées et considérées comme portant sur la mécanique classique sous toutes ses formulations. Il ajoute en effet une note particulièrement explicite à propos de l'existence de plusieurs formulations distinctes de la mécanique :

Il convient de dire quelques mots du choix de la formulation newtonienne comme base de la présente discussion. Il existe, certes, d'autres formulations de la théorie mécanique que celle de Newton, par exemple celles de Lagrange et de Hamilton. Ces formulations permettent d'analyser plusieurs problèmes difficiles avec plus de facilité et de souplesse que ce qui peut être fait au moyen de la formulation newtonienne. Cependant, ces autres formulations sont mathématiquement équivalentes au schéma newtonien, et l'on ne gagnerait rien à prendre comme point de départ l'une d'elles, d'autant plus qu'elles sont généralement moins familières. [...] Bien qu'un système de mécanique puisse rejeter la notion de force comme idée théorique première, et puisse même se dispenser entièrement d'utiliser le mot « force », il est toujours possible d'introduire le terme dans le système au moyen d'une définition nominale. Qui plus est, comme il apparaîtra bientôt, les différences entre les formulations de la mécanique [...] n'ont aucune conséquence sur l'analyse de la notion d'explication mécanique. (Nagel, 1961, p. 158, note 5)

Non seulement les différentes formulations de la mécanique sont équivalentes empiriquement – sinon pour la résolution pratique de problèmes – mais encore, et c'est là le plus surprenant à la lumière des remarques faites précédemment, elles n'affectent pas la notion d'explication mécanique. Qu'une formulation suggère l'existence de forces quand une autre semble parler d'énergie n'a aucune incidence, d'après ces quelques lignes, sur la manière dont elles expliquent les phénomènes du mouvement. Sans entreprendre d'analyser la manière dont Nagel rend compte de la notion d'explication mécanique, qui mériterait un examen détaillé, je conclurai simplement en suggérant que cette manière d'évacuer le problème posé par la multiplicité des formulations de

⁷⁸En ce sens, la méthode de Nagel n'est pas caractéristique, ici, de la « conception syntaxique » telle que les défenseurs de la conception sémantique la décrivent.

la mécanique est l'expression d'un attachement à la conception empiriste logique des théories, dont Nagel souligne pourtant les limites en introduisant dans sa définition des théories la notion de modèle. Cette incohérence manifeste dans l'analyse du cas de la mécanique classique illustre de manière assez remarquable la tension dont j'ai montré qu'elle caractérisait la conception nagélienne des théories.

1.7 Conclusion : la conception « syntaxique » des théories scientifiques

On a vu comment le projet formaliste, lié à l'exigence carnapienne de trouver un critère de signification cognitive, a conduit à une conception des théories scientifiques, que ses critiques ont appelé tantôt « conception orthodoxe » (Putnam, 1962), tantôt « conception syntaxique » ou « linguistique », et dont Suppe (1977a) a proposé une reconstruction extrêmement minutieuse (reproduite *in extenso* en appendice de ce chapitre, page 276), dont il reconnaît qu'elle ne peut être attribuée à aucun philosophe en particulier, n'ayant fait l'objet d'aucun « manifeste ». Les différents « piliers » sur lesquels elle repose (distinction théorie / observation, distinction analytique / synthétique, entre autres) ont fait l'objet de très nombreux débats au sein de la tradition empiriste logique⁷⁹ ; outre Nagel, dont on voit qu'il se démarque assez nettement de l'image d'une théorie constituée de deux composantes principales (qui est au cœur de la reconstruction, par Suppe, du programme des empiristes logiques), Hempel lui-même, à qui on attribue traditionnellement la défense de la conception syntaxique au même titre qu'à Carnap, critique les positions de ce dernier sur un certain nombre de points, qui le conduisent à émettre de sérieux doutes sur l'utilité même de la formalisation des théories (Hempel, 1970).

Ce que l'on a baptisé « conception syntaxique » des théories et qui constitue la première des deux approches formelles étudiées dans le présent chapitre ne peut pas être décrit comme un projet de formalisation effective des théories existantes : il s'agit plutôt de la description d'une forme canonique idéale que devrait pouvoir revêtir, en droit, toute théorie scientifique. Carnap (1966) lui-même, le plus attaché aux grands principes du premier positivisme logique, propose finalement une image assez libérale des théories scientifiques et ne prétend pas que la présentation d'une théorie sous cette forme canonique ferait gagner beaucoup à la science.

La « caricature » de la conception syntaxique n'est pourtant pas montée de toutes pièces. Par delà leurs divergences, ce qui caractérise aussi bien Nagel ou Hempel que Carnap, c'est qu'ils restent tous attachés à une image des théories scientifiques comme

⁷⁹Pour une recension détaillée de ces débats, voir l'introduction de Suppe (1977a) à (Suppe, 1974/1977b).

celle d'un réseau de principes et de lois dont il s'agit de décrire le lien à l'observable sur le modèle de la représentation linguistique. On en trouve une expression métaphorique dans ces propos de Hempel :

Une théorie scientifique peut donc être comparée à un réseau spatial complexe : ses termes sont représentés par les nœuds du réseau, tandis que les fils qui relient ces derniers correspondent en partie aux définitions, en partie aux hypothèses primitives ou dérivées que renferme la théorie. Le système tout entier flotte, pour ainsi dire, au-dessus du plan d'observation où il se trouve ancré au moyen des règles d'interprétation. On peut comparer ces règles à des chaînes qui n'appartiennent pas au réseau mais en relient certaines parties à des régions déterminées du plan d'observation. Grâce à ces relations qui servent à interpréter, le réseau peut remplir sa fonction de théorie scientifique : étant donné certains faits d'observation, nous pouvons remonter, par l'intermédiaire d'une chaîne d'interprétation, à un point précis du réseau théorique et, de là, par l'intermédiaire des définitions et des hypothèses, nous diriger vers d'autres points, d'où une nouvelle chaîne d'interprétation nous permettra de regagner le plan d'observation. (Hempel, 1966)⁸⁰

Le schéma de Feigl (1970), reproduit ci-dessous (figure 10), exprime cette idée d'un lien direct entre les définitions, lois et hypothèses d'un côté, et l'observable de l'autre. C'est cette image qui conduit les empiristes logiques à penser que, même

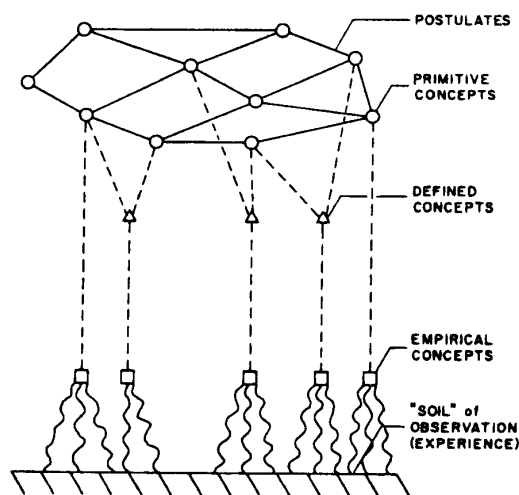


FIG. 10 – Schéma de la théorie scientifique selon l'empirisme logique (Feigl, 1970, reproduit par Giere, 1988, p. 25)

s'il est strictement impossible et inutile de la formaliser ainsi, la structure d'une théorie doit être pensée sur le modèle d'une syntaxe non interprétée, directement

⁸⁰ Je cite la traduction française indiquée en bibliographie.

reliée à l'observable. C'est l'attachement à cette image, joint à la reconnaissance du fait qu'une théorie est, de fait, toujours déjà interprétée, qui conduit Nagel à l'idée intenable d'une double sémantique.

La conception sémantique va consister à rejeter en bloc cette image « linguistique » des théories : prenant acte du fait qu'on ne peut pas séparer le squelette syntaxique d'une théorie de son interprétation, elle va chercher du côté de la sémantique formelle et de la théorie des modèles les outils nécessaires à la reconstruction du contenu des théories, qui n'est plus identifié à leur signification empirique. Ainsi, comme on va le voir, la signification d'une théorie ne lui est pas fournie par des règles de correspondance qui viennent, *après coup*, fournir une signification empirique à des axiomes non interprétés ; une théorie a, *d'emblée*, une interprétation, consistant en un ensemble de structures. Tout en renonçant au modèle de la traduction pour adopter celui de l'interprétation (au sens de la théorie des modèles), la conception sémantique reste cependant une approche formelle : son but est de rendre compte du contenu des théories d'une manière qui permette d'évacuer les aspects liés à sa (ou à ses) formulation(s) linguistique(s) particulière(s) et à son (leur) utilisation par les agents. Procéder à son examen va me permettre de mettre en évidence les limites de toute entreprise de ce type – et non pas seulement de celle des empiristes logiques.

2 La conception sémantique des théories scientifiques

Au cours des années 1960, une « nouvelle orthodoxie » (Le Bihan, 2006) en philosophie des sciences s'impose, incarnée par une autre approche formelle des théories, la « conception sémantique », présentée comme une critique des grands principes de la philosophie des sciences des empiristes logiques, qu'elle fige en un corps de doctrine baptisé « conception syntaxique »⁸¹ pour se définir par contraste avec elle.

Avant de faire l'objet de véritables manifestes qui l'instituent explicitement en nouvelle théorie des théories, elle voit le jour à la fin de la décennie 1940 avec les travaux du logicien néerlandais Evert Beth (1940), puis, au cours de la décennie 1950 avec ceux de Patrick Suppes (1957, 1960, 1962, 1967), qui voit dans la théorie des ensembles et la notion de modèle en sémantique formelle des outils prometteurs pour axiomatiser les théories scientifiques.

Outre Suppes, père et défenseur de cette approche, les grands noms en sont Frede-

⁸¹ Comme on va le voir, certains traits de l'image que les défenseurs de la conception sémantique proposent du programme des empiristes logiques sont contestables ; dans ce qui suit, l'expression « conception syntaxique » me servira par conséquent à désigner, plutôt que le projet empiriste logique tel que je le comprends et que je l'ai présenté dans ce qui précède, le portrait qu'en dresse la conception sémantique.

rick Suppe (1971, 1977a, 1989), Bas van Fraassen (1980, 1987, 1989, 1991) et Ronald Giere (1979, 1988, 2006). Une autre école qui s'en distingue plus nettement mais qui se situe malgré tout dans l'héritage de Suppes est l'école dite « structuraliste » de Wolfgang Balzer (1985), Ulises Moulines (1975), Joseph Sneed (1975, 1976) et Wolfgang Stegmüller (1976).⁸² Enfin, l'axiomatisation de plusieurs théories existantes suivant les méthodes de la conception sémantique a été proposée, entre autres, en physique (Hughes, 1989; van Fraassen, 1991) et en biologie (Beckner, 1959; Beatty, 1982; Lloyd, 1988; Thompson, 1989, 2007).

On résume généralement le « dogme central » de la conception sémantique comme suit : une théorie scientifique doit être décrite comme une classe (ou famille) de modèles plutôt que comme un ensemble d'énoncés. On ajoute généralement qu'une théorie est une entité extra-linguistique qui ne peut être identifiée à aucune de ses formulations particulières mais correspond à la classe des modèles qui satisfont ses différentes formulations linguistiques. La notion de modèle est ici à comprendre dans le sens logique qu'elle a dans le cadre de la théorie des modèles issue des travaux de Tarski en sémantique formelle : c'est une structure qui satisfait une théorie (voir ci-dessus, page 190).

Les défenseurs de cette approche affirment sa supériorité sur celle des empiristes logiques en plusieurs points. En premier lieu, elle ne se laisse pas égarer par des questions qui, comme celle de la signification des termes théoriques, sont certes des problèmes importants de philosophie du langage, mais n'ont aucune place, selon eux, dans une réflexion sur les sciences.⁸³ Ensuite, elle est censée mieux rendre compte de la pratique scientifique, qui consiste en la construction et la manipulation de modèles bien plus qu'en la déduction de descriptions phénoménales à partir de lois théoriques. Enfin, elle permet de réellement explorer le contenu de théories existantes par leur axiomatisation effective, alors que la formalisation rêvée par la conception syntaxique non seulement n'est pas utilisable en pratique mais encore, à supposer qu'elle le soit, ne permet pas d'éclairer certains aspects des théories et de leurs relations à d'autres théories mis en évidence par l'approche sémantique.

Dans cette section, je propose de présenter, principalement à partir de certains textes de Patrick Suppes⁸⁴, une analyse critique de la manière dont cette approche en-

⁸²Voir aussi (Balzer *et al.*, 1987).

⁸³ « Les distinctions scolastiques logicistes que la tradition positiviste logique a produit – vocabulaire observationnel et théorique, réductions de Craig, énoncés de Ramsey, théories axiomatisables en logique de premier ordre, et aussi prédicats projetables, énoncés de réduction, termes dispositionnels, et tout le satané reste – nous a conduit à *mille milles de toute habitation scientifique* [en français dans le texte], isolés dans nos rêves abstraits. » (van Fraassen, 1987, p. 110)

⁸⁴La thèse de doctorat de Patrick Suppes, sur le concept d'action à distance dans les travaux de

tend reconstruire le contenu des théories. J'expliquerai d'abord ce qu'il faut entendre par l'affirmation selon laquelle une théorie est une famille de modèles (section 2.1) ; comme on le verra, cette affirmation, qui est sans doute la plus unanimement partagée par tous les défenseurs de cette approche, consiste en un refus de la conviction des empiristes logiques selon laquelle la signification d'une théorie lui est fournie par son interprétation empirique. La section suivante (section 2.2) est consacrée à l'analyse de la manière dont les théories, ainsi conçues, reçoivent une interprétation empirique. Dans un troisième temps (section 2.3), j'examinerai les conséquences de cette distinction entre le contenu d'une théorie et sa signification empirique pour la question du statut cognitif des théories. Enfin (section 2.4), je tenterai de préciser les buts que cette conception assigne à la reconstruction formelle des théories, afin d'en évaluer le succès et d'en tracer les limites. Cela me permettra, dans la section 3, de revenir sur la distinction entre formulation et reconstruction formelle à la lumière de l'exemple de la mécanique classique ; j'y montrerai alors que, quand bien même la conception sémantique s'est construite sur un constat de l'échec de l'entreprise des empiristes logiques, elle ne parvient pas entièrement à dégager le contenu des théories des aspects, jugés contingents, liés à leurs formulations particulières et à leur utilisation par les agents.

2.1 Les théories comme familles de modèles

Que signifie l'affirmation selon laquelle on peut identifier le contenu d'une théorie à une classe de modèles ? Rappelons l'image des théories donnée par Nagel. Reconnaissant d'une part qu'une théorie, si on l'identifie à un squelette syntaxique et à des règles reliant les énoncés non interprétés de ce squelette à des énoncés doués d'une signification empirique, n'a pas de réel contenu sémantique – elle ne dit rien –, mais restant d'autre part attaché à l'idée d'une théorie comme ayant en son cœur une structure syntaxique, il est conduit à donner une image hybride intenable de la sémantique des théories. La conception sémantique fait un pas de plus. Prenant acte du fait, mis en évidence par Nagel, qu'une théorie est toujours déjà interprétée, elle refuse purement et simplement les deux postulats suivants : celui selon lequel on peut (au moins en principe) isoler le contenu de la théorie en reconstruisant son squelette syntaxique, et celui selon lequel c'est par l'établissement d'une relation linguistique entre ce squelette et des énoncés empiriques que la théorie reçoit sa signification.

Une fois les règles de correspondance écartées et le divorce entre les questions de

Descartes, Newton, Boscovitch et Kant, était dirigée par Ernest Nagel. Merci à Paul Humphreys (qui a lui-même écrit sa thèse de doctorat sous la direction de Patrick Suppes) pour cette indication. Voir l'« autobiographie intellectuelle » de Patrick Suppes dans (Bogdan, 1979).

l'application empirique et du contenu sémantique prononcé, ce dernier est entièrement donné par la classe des modèles que la théorie décrit.⁸⁵ La notion de modèle, beaucoup plus nettement que chez Nagel, est ici à comprendre dans le sens qu'elle a en logique mathématique : un modèle est une structure qui satisfait un énoncé ou un ensemble d'énoncés.⁸⁶ En conséquence, le contenu d'une théorie, selon la conception sémantique, est donné par *sa théorie mathématique*, c'est-à-dire l'ensemble des modèles qui en satisfont les différentes formulations.

Afin de mieux comprendre ce que cela signifie, je propose de présenter brièvement la méthode d'axiomatisation mise en œuvre dès les années 1950 par Patrick Suppes (1957). Comme je l'ai dit, la conception sémantique consiste, plus qu'en une théorie philosophique des théories, en un programme effectif d'axiomatisation des théories scientifiques existantes. Suppes propose ainsi l'axiomatisation de plusieurs théories, dont la mécanique classique du point⁸⁷, selon une méthode qu'il juge plus efficace que celle des axiomatisations syntaxiques ; il admire la rigueur de la démarche axiomatique et pense qu'elle est susceptible, une fois armée des bons outils, de clarifier plusieurs problèmes philosophiques et scientifiques (voir Suppes, 1968 et, ci-dessous, section 3.1 du présent chapitre). Cette méthode, inspirée de certaines idées de Tarski et des travaux sur les fondements des mathématiques entrepris, dans la lignée de Hilbert (1899), par le groupe Bourbaki pendant les années 1940 et 1950, consiste à définir un prédicat dans les termes de la théorie des ensembles (un *prédicat ensembliste*). Toute entité qui satisfait un tel prédicat – de laquelle ce prédicat est vrai – est un modèle de la théorie, au sens logique du terme modèle (Suppes, 1957, pp. 249-253). L'hypothèse de Suppes est que cette méthode, jusque là appliquée aux théories mathématiques, peut servir à reconstruire le contenu des théories *empiriques*.

⁸⁵La sémantique est toujours définie en termes de vérité. Seulement, au vérificationnisme qui conçoit la signification d'un énoncé comme donnée par ses conditions empiriques de vérification, on substitue la notion de vérité *dans un modèle*. La signification des énoncés d'une théorie n'est plus donnée par l'état du monde qui les rendrait vrais, mais par l'ensemble des structures qui les satisfont (dans lesquelles ils sont vrais).

⁸⁶Notons bien la chose suivante : les modèles, chez Nagel, sont censés fournir une interprétation à la théorie, ce qui encourage à comprendre le terme en son sens logique ; cependant, on l'a vu, un modèle fournit aussi une analogie, suggérée par l'emploi d'un certain type d'équations ou plus largement d'un certain langage, et est à ce titre attaché à une certaine formulation. On ne peut donc pas décrire l'usage de la notion de modèle par la conception sémantique comme un prolongement pur et simple de celui qu'en fait Nagel. Je reviendrai à la fin de ce chapitre (page 274) sur cet important problème.

⁸⁷McKinsey, Sugar et Suppes (1953) avaient déjà proposé une axiomatisation de la mécanique classique, que Suppes reprend et développe en une conception des théories scientifiques après la mort de McKinsey en 1953. Son *Introduction to Logic* (1957) est dédiée à la mémoire de ce dernier, qui était, précisons-le, un disciple de Tarski (voir Moulines, 2006, p. 110).

Un prédicat ensembliste permet de définir un ensemble de structures au moyen d'une suite de la forme $\langle D_1, \dots, D_m, R_1, \dots, R_n \rangle$, où les D_i représentent les « domaines de base » de la théorie (par exemple, un ensemble fini d'objets, un continuum, etc.) et les R_i les relations construites sur ces domaines (par exemple, une relation symétrique et transitive, une fonction deux fois différentiable sur les nombres réels, etc.). Une fois cette suite définie, on lui impose un certain nombre de contraintes en formulant les *axiomes* de la théorie. Prenons l'exemple de l'axiomatisation de la mécanique newtonienne du point proposée par Suppes (1957).

Après s'être donné les notions d'« intervalle de nombres réels » de « vecteur » de « différentiabilité d'une fonction », et défini P et T comme des ensembles, s comme une fonction binaire, m comme une fonction unaire, t comme une fonction ternaire et g une fonction binaire, voici comment il définit le prédicat ensembliste que doit satisfaire tout système de la mécanique du point (Suppes, 1957, p. 294) :

DEFINITION 1 *Un système $\beta = \langle P, T, s, m, f, g \rangle$ est un système de la mécanique du point si et seulement si les sept axiomes suivants sont satisfaits :*

AXIOMES CINÉMATIQUES

AXIOME P1. *L'ensemble P est fini et non vide.*

AXIOME P2. *L'ensemble T est un intervalle de nombres réels.*

AXIOME P3. *Pour p dans P , s_p est deux fois différentiable par rapport à T .*

AXIOMES DYNAMIQUES

AXIOME P4. *Pour p dans P , $m(p)$ est un nombre réel positif.*

AXIOME P5. *Pour p et q dans P et t dans T ,*

$$f(p, q, t) = -f(q, p, t).$$

AXIOME P6. *Pour p et q dans P et t dans T ,*

$$s(p, t) \times f(p, q, t) = -s(q, t) \times f(q, p, t).$$

AXIOME P7. *Pour p dans P et t dans T ,*

$$m(p) \frac{d^2 s_p}{dt^2}(t) = \sum_{q \in P} f(p, q, t) + g(p, t)$$

Les axiomes ci-dessus donnent la définition d'un système de la mécanique du point : toute structure, abstraite ou concrète, satisfaisant la description cinématique et dynamique du système β (tout modèle satisfaisant les axiomes $P1$ à $P7$) est un système de la mécanique du point. Le contenu de la mécanique classique du point est constitué de l'ensemble de ces systèmes. La classe de toutes les structures (modèles logiques) qui la satisfont est ce que Suppes propose d'appeler le « modèle mathématique » de la théorie. Par extension, on appelle parfois « modèle » l'équation qui décrit ce modèle (Suppes, 1957, pp. 253-254).

Ces axiomes ne sont donc pas des axiomes du même type que les postulats théorique (T) des empiristes logiques : ce ne sont pas des *énoncés non interprétés* exprimant des lois⁸⁸ – ou plutôt des formes de lois – auxquelles il s'agirait d'attribuer des référents dans le monde empirique pour pouvoir en déduire des énoncés empiriques et, ce faisant, leur donner un contenu. Pour reprendre la distinction signalée page 192, cette axiomatisation n'est pas une *formalisation au sens strict*. En effet, le prédicat ensembliste est défini au moyen de tout l'appareil mathématique jugé utile à l'expression de la théorie, sans qu'il soit nécessaire de fonder, à son tour, cet appareil mathématique au moyen d'une formalisation stricte. La méthode de Suppes, comme celle des axiomatisations modernes de la géométrie dont elle est inspirée, ne consiste pas à *mettre en relation une syntaxe et une sémantique, conçue sur le modèle de la vérifiabilité empirique*, mais à se donner un ensemble d'éléments et des objets mathématiques (relations, fonctions, opérations sur cet ensemble), puis à les soumettre à des conditions énoncées par les axiomes.⁸⁹

Prenons l'axiome $P7$: il correspond à l'expression du principe fondamental de la dynamique. Une expression purement syntaxique de cette loi ne consisterait qu'en un énoncé établissant une certaine relation entre deux variables F et a , sans rien dire de la manière dont il convient d'interpréter ces termes (indiquer que ce sont des vecteurs, c'est déjà en donner une interprétation, sinon physique, du moins mathématique). L'interprétation empirique de ces termes serait assurée par les règles de correspondance qui spécifieraient les opérations de mesure permettant d'attribuer une valeur à ces variables dans le cadre de l'étude du comportement d'un système empirique. Ici, en revanche, l'axiome $P7$, s'il ne donne encore aucun moyen de décrire un système empirique, décrit un objet mathématique : une classe de structures, définies dans un certain domaine, qui obéissent à certaines conditions.

Les axiomes utilisés pour présenter la théorie, dans la perspective sémantique, sont donc des axiomes déjà interprétés, sinon physiquement, du moins mathématiquement. Les modèles ainsi définis ne sont pas les modèles d'une structure syntaxique

⁸⁸La notion de « loi » est une des notions traditionnelles de la philosophie des sciences qui est le plus remise en question par la conception sémantique des théories. Ainsi Giere (1988) intitule-t-il son ouvrage *Science without laws*, quand van Fraassen (1989) propose le remplacement de cette notion par celle de symétrie. Enfin, Lloyd (1988), à la suite de Beatty (1982), montre que la notion de loi qui accompagne l'image hypothético-déductive de la science véhiculée par l'empirisme logique est particulièrement mal adaptée à l'étude de la biologie évolutionniste : la loi de Hardy-Weinberg, par exemple, ne correspond pas à ce que les empiristes logiques appellent une « loi ».

⁸⁹« Dans une présentation moderne de la géométrie, on ne trouve pas les axiomes de la géométrie euclidienne, mais la définition d'un espace euclidien. De même, Suppes et ses collaborateurs ont cherché à reformuler les fondements de la mécanique newtonienne en remplaçant les axiomes de Newton par la définition d'un système mécanique newtonien. » (van Fraassen, 1987, p. 109)

non interprétée, mais des objets mathématiques décrits par un langage qui se donne tout l'appareil mathématique nécessaire à la présentation directe de la structure parfois mathématiquement très sophistiquée des théories physiques les plus complexes. Le grand pas franchi par la conception sémantique consiste à nier que l'exigence fondationnaliste qui guide la démarche des empiristes logiques ait la moindre pertinence pour l'étude des théories scientifiques. Si le projet logiciste d'étude des fondements des mathématiques est de toute première importance pour la philosophie des mathématiques et la logique – « la métamathématique », comme l'appelle Suppes –, l'ambition d'une axiomatisation des théories des sciences empiriques à partir d'une syntaxe non interprétée est, selon lui, un projet irréalisable et vain.

Irréalisable, parce que beaucoup de théories scientifiques, comme la mécanique quantique ou la thermodynamique classique, par exemple, qui supposent qu'on utilise « non seulement les idées générales de la théorie des ensembles, mais également plusieurs résultats concernant les nombres réels », ont une structure d'un degré de complexité semblable à celui des théories étudiées dans les mathématiques pures, dont la formalisation stricte est pratiquement impossible. « Dans de tels contextes, il est beaucoup plus simple d'affirmer des choses sur les modèles de la théorie que de parler directement et explicitement des énoncés de la théorie, peut-être en tout premier lieu parce que la notion d'un énoncé de la théorie n'est pas bien définie quand la théorie n'a pas de formalisation standard » (Suppes, 1967, p. 58). Ici, par « formalisation standard », Suppes entend « formalisation en logique du premier ordre ». Comme je l'ai signalé au début de ce chapitre (note 7), un des reproches que la conception sémantique fait au projet des empiristes logiques est en effet d'exiger une formalisation des théories dans un langage du premier ordre. Or une telle formalisation est souvent impossible. La définition de plusieurs concepts mathématiques requiert en effet une logique d'ordre supérieur. Ce reproche, comme je l'ai dit, n'est pas historiquement fondé, au sens où Carnap n'a jamais, à ma connaissance, formulé une telle exigence.⁹⁰ On peut cependant interpréter ces propos de Suppes comme une rationalisation *a posteriori* du projet des empiristes logiques : formaliser en logique de premier ordre rendrait leur approche parfaitement cohérente, dans la mesure où les langages de premier ordre sont des langages complets.⁹¹

Je propose cependant de comprendre cette caractérisation de la conception syntaxique – et son rejet – comme l'expression non pas tant du *choix d'outils formels*

⁹⁰ Il admet en effet (voir ci-dessus, note 69 page 226) que l'on peut présenter une théorie à l'aide de tout l'appareil mathématique qu'on voudra ; il ne refuse pas l'usage des mathématiques dans la reconstruction formelle.

⁹¹ Dans un langage de premier ordre, quand on a une implication, on a aussi une conséquence logique dans tous ses modèles, ce qui n'est pas le cas dans un langage d'ordre supérieur.

différents (théorie des ensembles contre logique du premier ordre)⁹², que comme celle d'une *conception différente du but de l'entreprise axiomatique elle-même* : cette dernière n'est pas destinée à analyser le langage des théories, mais les structures mathématiques décrites au moyen de ce (ou ces) langage(s), quel(s) qu'il(s) soi(en)t. Autrement dit, le problème n'est pas tant de savoir s'il est possible ou non de trouver un langage logique (de premier ordre ou d'ordre supérieur) permettant de formaliser strictement les théories mathématiques utilisées dans la reconstruction formelle des théories empiriques, que celui de la pertinence même d'une telle entreprise.⁹³

C'est en ce sens que le projet de la « conception syntaxique », à supposer même qu'il soit réalisable, est *vain*. Formaliser strictement les portions des théories mathématiques dont on a besoin pour exprimer une théorie scientifique serait une tâche inutilement laborieuse, comme le serait « l'axiomatisation simultanée [de] la géométrie *et* [des] portions pertinentes de la théorie des ensembles » nécessaires à la définition, par exemple, des droites « comme certains *ensembles* de points » (Suppes, 1967, p. 58). L'axiomatisation syntaxique des théories revient ainsi à se priver délibérément d'outils indispensables à l'exploration de la structure des théories, en raison d'une précaution reposant sur une erreur : l'idée selon laquelle une théorie n'a pas de signification si elle n'a pas de signification empirique. Cette idée entraîne en effet

⁹²D'ailleurs, plusieurs défenseurs de la conception sémantique proposent des manières différentes de concevoir et de présenter les modèles. L'école structuraliste de Sneed et Stegmüller suit Suppes en les décrivant comme des entités ensemblistes, tout en ajoutant d'autres composantes indispensables, selon eux, à l'identification d'une théorie (voir Balzer *et al.*, 1987; Moulines, 2002, 2006) ; Bas van Fraassen, inspiré par Evert Beth, et Frederick Suppe proposent indépendamment l'un de l'autre de les concevoir comme des trajectoires dans des espaces des phases ; c'est aussi ce dernier type d'axiomatisation que les défenseurs de la conception sémantique en philosophie de la biologie comme Lloyd et Thompson ont tendance à utiliser. Malgré ces différences dans la manière de mettre en œuvre l'axiomatisation, il existe une communauté fondamentale de vue : « Bien que je ne souhaite élire aucune présentation mathématique particulière au rang de méthode canonique, j'adopte clairement [la] conception générale [de Suppes] de la manière dont la mécanique classique doit, disons, être identifiée » (van Fraassen, 1980, p. 66)

⁹³Comme me l'a fait remarquer Paul Humphreys dans une communication personnelle, on pourrait envisager une formalisation syntaxique d'une théorie comme la mécanique quantique, qui utilise la théorie des groupes, en donnant une définition d'un groupe (clôture sous une opération donnée, existence d'une opération inverse, existence d'un élément d'identité, satisfaction de l'associativité) qui permettrait de le considérer comme un objet purement formel. Autrement dit, si le but n'est pas tant de reformuler une théorie dans un langage donné que de rendre compte des opérations logiques elles-mêmes, il n'est pas certain qu'une axiomatisation strictement syntaxique soit impossible. La plupart des philosophes des mathématiques nieraient cependant la possibilité d'un tel formalisme pur. Quoiqu'il en soit, dans les axiomatisations telles que Suppes les envisage, il est clair que la théorie des ensembles et tout un appareil mathématique que l'on ne peut pas formaliser strictement doivent déjà être disponibles.

les empiristes logiques à faire un détour forcé par l'analyse du langage des mathématiques alors que ce qui intéresse le philosophe des sciences, selon les défenseurs de la conception sémantique, est la structure mathématique elle-même.

Quand les théories sont formalisées en logique de premier ordre, les théorèmes qui relient les différents modèles de la théorie sont nécessairement métamathématiques dans la manière dont ils sont énoncés et prouvés.⁹⁴ Par contraste, les théorèmes comparant différents modèles d'une théorie peuvent être énoncés de façon d'emblée mathématique, quand la théorie est axiomatisée par la définition d'un prédicat ensembliste. (Suppes, 1957, p. 254)

Tout comme le mathématicien qui étudie des structures d'une grande complexité, le logicien philosophe a tout intérêt à se donner un langage aussi riche que nécessaire ; plutôt que de « repasser » systématiquement au niveau métamathématique, il peut ainsi considérer directement les structures mathématiques étudiées et leurs relations : les outils de la théorie des modèles, définissant différentes relations structurelles entre modèles (isomorphisme, plongement)⁹⁵ permettent ainsi de résoudre beaucoup plus efficacement que ne le permet une étude du langage des théories certains problèmes traditionnels de la philosophie des sciences. Par exemple, la notion de réduction inter-théorique, traitée de manière approfondie par Nagel (1961, chap. 11) dans la tradition syntaxique, est selon Suppes très efficacement étudiée par l'établissement d'un théorème de représentation⁹⁶ entre les modèles de la théorie réduite et ceux de la théorie réductrice : « la thèse selon laquelle la biologie peut être réduite à la physique serait établie de façon appropriée dans beaucoup d'esprits si l'on pouvait montrer que pour tout modèle d'une théorie biologique il est possible de construire un modèle isomorphe au sein de la théorie physique » (Suppes, 1960, p. 295).⁹⁷

⁹⁴[note de Suppes] : « Généralement, les modèles métamathématiques ne sont pas utilisés par les mathématiciens quand ils peuvent être évités, car leur application exacte exige que l'on travaille avec un langage entièrement spécifié et formalisé, ce qui est une tâche souvent fastidieuse et difficile. Cette remarque n'est pas destinée à dévaloriser de quelque manière que ce soit les métamathématiques. Plusieurs résultats importants ne peuvent être établis que par le biais de méthodes métamathématiques. »

⁹⁵Un isomorphisme est, littéralement, une identité de structure. Deux modèles sont isomorphes s'il existe entre eux une application bijective qui fait correspondre leurs constantes et leurs relations de façon univoque. Un plongement (*embedding*) est une relation structurelle moins forte : sans entrer dans plus de détails techniques, disons que l'on parle de plongement quand l'un des modèles est isomorphe à une partie (une sous-structure) de l'autre (voir Suppes, 1988).

⁹⁶« Etablir un théorème de représentation pour une théorie, c'est prouver qu'il y a une classe de modèles de la théorie telle que chaque modèle de la théorie est isomorphe à un membre de cette classe. » (Suppes, 1960, p. 295).

⁹⁷On peut aussi établir un théorème de représentation entre deux branches de la physique : Suppes mentionne le résultat d'Adams (1959) qui établit que tout modèle de la mécanique des solides est

Le projet axiomatique en philosophie des sciences, selon la conception sémantique, a donc pour but de nous permettre d'explorer directement la structure mathématique des théories, et non de fonder les mathématiques (ou de s'en passer). C'est ce que signifie le slogan de Suppes, rapporté par Bas van Fraassen (1980, p. 65), selon lequel « *la philosophie des sciences doit utiliser les mathématiques, et non la métamathématique* ». La conception sémantique remplace ainsi la logique et l'analyse du langage par les mathématiques et l'exploration des structures comme outils pour étudier le contenu des théories. Semblant simplement proposer de nouveaux outils pour axiomatiser les théories scientifiques dans la veine des empiristes logiques, c'est en fait à un rejet complet de leur programme qui consiste à fonder la signification des théories sur leur vérifiabilité empirique que Suppes procède. Pour lui, axiomatiser une théorie scientifique, c'est exprimer – en s'autorisant tous les outils nécessaires – *sa théorie mathématique*, c'est-à-dire la classe des modèles satisfaisant ses différentes formulations linguistiques.⁹⁸

L'exigence des empiristes logiques reposait sur la croyance erronée selon laquelle les énoncés utilisés pour exprimer une théorie acquièrent leur signification de leur lien à l'observable. Pour la conception sémantique, le langage utilisé pour axiomatiser les théories sert à décrire des structures mathématiques ; il n'a pas à être relié directement aux phénomènes empiriques. Une fois ce divorce prononcé entre la question de l'interprétation et celle de l'application empirique, comment les théories sont-elles reliées au monde et comment acquièrent-elles un contenu empirique ? Qu'est-ce qui fait d'une théorie scientifique ainsi définie une théorie *empirique* ? La réponse de la conception sémantique est que ce sont les modèles eux-mêmes – et non leur description linguistique – qui sont mis en relation avec les phénomènes empiriques. Comme on va le voir, cette analyse prétend rendre compte bien plus fidèlement que ne le permet la conception syntaxique de la réalité de l'activité théorique.

isomorphe à un modèle défini au sein de la mécanique du point (Suppes, 1960, pp. 295-296). Pour une étude détaillée de l'établissement de théorèmes de représentation entre (et au sein de) théories issues de différents domaines, voir (Suppes, 2002).

⁹⁸ Il convient d'insister : les *formulations linguistiques* dont il s'agit, contrairement à ce que laissent trop souvent entendre les avocats ou les exégètes de la conception sémantique (voir page 265), ne sont pas les formalisations strictes proposées – ou plutôt rêvées – par les empiristes logiques. Même si sur ce point les différents exposés ou manifestes de la conception sémantique sont loin d'avoir le vocabulaire unifié qui permettrait de clarifier leur propos, je ne vois pas quel sens on pourrait donner à l'affirmation selon laquelle les modèles qui constituent le contenu de la théorie sont ceux qui en satisfont la formalisation stricte, puisque le principe même d'une telle formalisation est exclu dans un grand nombre de cas.

2.2 L'application empirique des théories : modèles logiques et modèles physiques

Tout autant que l'idée de la possibilité de dégager le squelette syntaxique des théories, celle de l'établissement de règles linguistiques permettant la mise en relation d'énoncés théoriques avec des énoncés observationnels repose, selon les défenseurs de la conception sémantique, sur une idéalisation abusive et stérile de la manière dont les théories sont appliquées à la prédiction et à l'explication des phénomènes empiriques. Un rapide examen de la pratique scientifique met en évidence le fait que les théories sont enseignées, développées et appliquées par la manipulation de modèles bien plus que par des raisonnements inductifs et déductifs sur des énoncés (comme on l'a vu à la section 1.3 du chapitre 2).

La notion de « modèle » désigne ici des choses de nature assez hétérogène. Des systèmes idéalisés comme le pendule simple en mécanique, l'agent parfaitement rationnel en économie, la population isolée en génétique des populations, mais également les équations ou les ensembles d'équations qui décrivent le comportement de ces systèmes, le modèle réduit d'un avion, la représentation concrète de la molécule d'ADN construite par Crick et Watson, le modèle – parfois appelé « théorie » – de l'atome de Bohr, sont tour à tour appelés « modèles ».

Pourtant, les modèles dont il était question à la section précédente sont des « entités hautement abstraites, non linguistiques, souvent très éloignées dans leur conception des observations empiriques » (Suppes, 1967, p. 57). Comment ces entités hautement abstraites sont-elles reliées aux phénomènes observables, et quel rapport y a-t-il entre elles et les modèles dont il est question au paragraphe précédent ?

C'est ici que le postulat le plus fort – et vivement critiqué au cours de ces dernières décennies⁹⁹ – de la conception sémantique intervient : non seulement les modèles théoriques, c'est-à-dire les structures décrites par la théorie, peuvent être conçus comme des modèles logiques, mais c'est aussi, selon Suppes, le cas des modèles que les scientifiques utilisent pour représenter les phénomènes, les prédire et les expliquer.¹⁰⁰ Ainsi,

⁹⁹Voir notamment (Frigg, 2006; Suárez, 1999, 2003; Thomson-Jones, 2006).

¹⁰⁰Si la notion logique de modèle désigne une structure particulière, la notion physique de modèle, comme celle de l'atome de Bohr, ou encore celle du pendule simple, désigne plutôt une classe de structures isomorphes. Mis à part cela, la différence entre les deux usages du terme « modèle » est moins importante qu'il n'y paraît : « Les scientifiques aussi parlent de modèles, et même de modèles d'une théorie, et leur usage est légèrement différent. "Le modèle de l'atome de Bohr", par exemple, ne fait pas référence à une unique structure. Cette expression fait plutôt référence à un type de structure, ou une classe de structures, qui partagent toutes certaines caractéristiques. En effet, selon cet usage, le modèle de Bohr devait convenir pour les atomes d'hydrogène, les atomes d'hélium, etc. Ainsi, dans l'usage que les scientifiques font du terme "modèle", il dénote ce que j'appellerai un type de modèle. À chaque fois que certains paramètres sont laissés sans spécification dans la description

après avoir cité des passages extraits aussi bien de textes de logique mathématique que de sciences empiriques, Suppes conclut :

Je soutiens que le concept de modèle dans le sens de Tarski peut être utilisé sans distorsion et comme un concept fondamental dans toutes les disciplines d'où les citations précédentes sont tirées. En effet, j'affirme que la signification du concept de modèle est la même en mathématiques et dans les sciences empiriques. La différence que l'on peut trouver dans ces disciplines est à chercher dans leur usage du concept. (Suppes, 1960, p. 12)

Il suffit selon Suppes de considérer que le modèle physique contribue à spécifier le domaine de discours du modèle plus abstrait, permettant par là de lui donner une signification empirique :

Définir formellement un modèle comme une entité ensembliste qui est un certain type de liste ordonnée consistant en un ensemble d'objets et de relations et d'opérations sur ces objets ne revient pas à écarter le modèle physique du type de ceux utilisés par les physiciens, car le modèle physique peut être simplement pris comme définissant l'ensemble des objets dans le modèle ensembliste. (Suppes, 1960, p. 13)

Pour prendre un exemple simple, le principe fondamental de la dynamique établit une relation de proportionnalité entre deux vecteurs \mathbf{F} et \mathbf{a} , le coefficient de proportionnalité étant appelé m ; il définit par là un certain type de structures mathématiques ; l'interprétation physique la plus générale de cette équation indique que \mathbf{F} est une force. On peut ensuite spécifier le type de force (électromagnétique, gravitationnelle, force de rappel...), ce qui permet de décrire le comportement d'un certain type de systèmes idéalisés (par exemple le pendule simple) ; on peut enfin attribuer une valeur précise à \mathbf{F} dans le cadre de l'étude du comportement dynamique d'un système particulier (le mouvement du balancier de telle horloge normande). On a donc, tout en bas de la hiérarchie, une certaine équation qui décrit adéquatement le comportement d'un système physique réel – que le comportement d'un système physique réel satisfait. À chaque étape, on a bien la description d'un certain type de structure, de plus en plus spécifique.

Les modèles situés tout en bas de la hiérarchie sont appelés « modèles de données » par Suppes (1960, 1962, 1957). Ce n'est pas l'objet ou le processus physique lui-même (le mouvement du balancier de l'horloge normande) qui satisfait l'équation mathématique, car il n'est pas, tel quel, identifiable à une structure. En revanche, son comportement est tel qu'il peut être modélisé, représenté, comme une structure

d'une structure, il serait plus précis de dire (contrairement à ce qui est l'usage commun) que nous décrivons un type de structure. Cependant, les usages de "modèle" en métamathématiques et dans les sciences ne sont pas si éloignés que ce que l'on dit parfois. » (van Fraassen, 1980, p. 44)

satisfaisant cette équation, avec le niveau de précision souhaité. Suppes parle d'une « version hautement schématisée de l'expérience » (Suppes, 1960, p. 300).

Suppes (1960, 1962, 1967) affirme que l'activité théorique peut être adéquatement décrite, à l'aide des outils formels de la théorie des modèles, comme consistant à établir une certaine relation fonctionnelle entre les structures (voir page 247)¹⁰¹ correspondant à des niveaux d'abstraction différents.¹⁰² À chacun de ces niveaux, une théorie correspondante décrit ce qu'en serait une réalisation possible : tout en haut de la hiérarchie, la théorie fondamentale décrit les modèles les plus abstraits ; puis, une théorie de l'expérience décrit les réalisations empiriques possibles de la théorie ; ensuite une théorie statistique définit la manière dont on peut construire les modèles de données à partir de l'expérience concrète.¹⁰³ Lors de ce qu'on appelle couramment « l'application d'une théorie à l'expérience », ce qui est comparé n'est pas la théorie abstraite d'un côté et les données de l'autre, mais les modèles de l'expérience et les modèles des données qui en sont des réalisations concrètes (Suppes, 1962, p. 253). De ce point de vue, la notion de règle de correspondance des empiristes logiques, mis à part tous les autres problèmes qu'elle pose, est une simplification abusive ; si l'on veut parler de règles de correspondance, ce n'est pas entre les modèles les plus abstraits de la théorie et l'expérience qu'il faut les établir, mais entre le modèle de l'expérience et celui des données (Suppes, 1967, p. 62). Cela permet de rendre compte de la difficulté suivante propre aux théories empiriques : contrairement à ce qui est le cas dans le domaine des mathématiques pures, les modèles qui sont comparés dans le processus de mise en relation de la théorie à l'expérience ne sont pas du même type logique ; les modèles de la théorie contiennent souvent des fonctions continues ou des séquences infinies, alors que ceux de l'expérience sont discrets et finis (Suppes, 1960, p. 297, 1962, p. 253, 1967, pp. 62-63).

Cette idée d'une hiérarchie de théories auxquelles correspondent autant de modèles de niveaux d'abstraction différents remet profondément en question l'idée, jugée naïve, de la confrontation d'une théorie à l'expérience : l'application empirique des

¹⁰¹Sur ce point aussi, des divergences de vues existent entre les différents tenants de la conception sémantique : van Fraassen et Suppe parlent d'isomorphisme, Giere parle de ressemblance. Ces différences, malgré leur importance, n'ont pas d'incidence sur la critique que je propose de la conception sémantique.

¹⁰²Par conséquent, l'établissement de théorèmes de représentation sert aussi bien à étudier les relations inter-théoriques que les relations structurelles sur lesquelles repose l'application empirique d'une théorie.

¹⁰³En dessous des modèles de données, on trouve encore ce que Suppes propose d'appeler la « méthode expérimentale » et les « conditions *ceteris paribus* », qui dirigent le montage expérimental lui-même (élimination du « bruit » et choix du dispositif expérimental.) Voir (Suppes, 1962, pp. 258-259).

théories est une « histoire beaucoup plus compliquée que les remarques habituelles sur les définitions de coordination et les interprétations empiriques des théories laissent penser. [...] une machinerie formelle plus élaborée et plus sophistiquée est requise pour relier la théorie aux données » (Suppes, 1967, p. 62). De la théorie fondamentale aux données brutes, l'interprétation empirique est affaire de degrés.

Un des péchés récurrents des philosophes des sciences est de simplifier outre mesure la structure de la science. Les philosophes qui écrivent sur la représentation des théories scientifiques comme des calculs logiques disent ensuite que la théorie reçoit une signification empirique par l'attribution de définitions interprétatives ou de coordination à certains des termes primitifs ou définis du calcul. Ce que j'ai tenté de montrer c'est que toute une hiérarchie de modèles se dresse entre le modèle de la théorie fondamentale et les résultats complexes de l'expérimentation. De plus, à chaque niveau de la hiérarchie il y a une théorie autonome [*in its own rights*]. Une théorie à un certain niveau reçoit une signification empirique par l'établissement de connexions formelles avec une théorie d'un niveau inférieur. L'étude statistique ou logique des relations entre théories à ces différents niveaux peut être faite d'une manière purement formelle, dans les termes de la théorie des ensembles. Plus l'analyse est explicite, moins il reste de place pour des considérations non formelles. Une fois les données empiriques mises sous forme canonique au niveau des modèles de données [...], toute question d'évaluation systématique qui se présente est une question formelle. (Suppes, 1962, pp. 260-261)

La méthode axiomatique de la conception sémantique se présente donc comme un moyen d'explorer systématiquement, au moyen d'outils formels, non seulement la structure mathématique des théories scientifiques, mais encore leur lien aux phénomènes empiriques. Comprendre une théorie, dans cette perspective, c'est donc, d'une part, être en mesure d'évaluer la relation entre les structures mathématiques qu'elle décrit et, d'autre part, savoir identifier, au sein de ces structures, les éléments qui ont un référent dans le monde empirique. On a vu, au cours de l'examen de la conception empiriste logique des théories, que la question de la signification des théories conduit à celle de leur statut cognitif, c'est-à-dire de leur prétention ou non à proposer une représentation vraie du monde. Dans la mesure où le contenu des théories, selon la conception sémantique, peut être étudié indépendamment de leur application empirique, il semble que la question de leur statut cognitif se pose très différemment : pour le dire grossièrement, il y a ce que la théorie dit « tout court » et ce qu'elle permet de dire à propos du monde empirique. Je voudrais à présent montrer, par un rapide examen des conséquences d'une telle distinction pour le problème du statut cognitif des théories, que ce problème ne peut pas être réglé au sein de la reconstruction formelle. Cela ne constitue pas, en tant que tel, une critique de cette conception (plu-

sieurs de ses défenseurs en conviennent). Ce sera cependant un élément déterminant de la critique que j'en proposerai dans la section 3 : on verra en effet que l'attitude épistémique que l'on adopte vis-à-vis d'une théorie intervient dans la reconstruction de son contenu.

2.3 Le statut cognitif des théories : que nous disent les théories scientifiques ?

Le fait de décrire le contenu sémantique des théories et le processus de leur application empirique à l'aide des outils formels de la théorie des modèles dissocie, comme on l'a vu, la question de l'interprétation de la théorie de celle de sa signification empirique et donc de sa vérité. La notion de vérité à l'œuvre jusqu'à maintenant est celle de vérité relative à un modèle : la théorie est vraie, par définition, des modèles qu'elle décrit. Comme le dit Ronald Giere,

La relation entre des équations (correctement interprétées) et leurs modèles peut être décrite comme une relation de caractérisation, ou même de définition. On peut même légitimement parler de « vérité » dans ce cas. Les équations interprétées sont *vraies* de leurs modèles. Mais la vérité, ici, n'a pas de signification *épistémologique*. Les équations décrivent de manière vraie le modèle parce que le modèle est défini comme quelque chose qui satisfait exactement les équations. (Giere, 1988, p. 79)

Cependant, comme le remarque Bas van Fraassen (1987, p. 106), une théorie « est le genre de chose que l'on peut accepter ou refuser, et croire ou ne pas croire [...]. Pour le dire plus généralement, une théorie est un objet d'attitudes épistémiques ou du moins doxastiques ». Autrement dit, une théorie a un contenu propositionnel, elle affirme quelque chose, qui peut être l'objet de différentes attitudes propositionnelles.¹⁰⁴

Rendre compte de cet aspect propositionnel des théories nécessite l'ajout d'une *hypothèse théorique*, celle qui est faite par l'utilisateur de la théorie lorsqu'il affirme l'existence d'une certaine relation d'isomorphisme ou de ressemblance¹⁰⁵ entre les modèles des différents niveaux de la hiérarchie. Cette hypothèse ne fait pas partie des axiomes qui décrivent un modèle ; si l'on entend par « théorie » uniquement la définition d'une classe de modèles, alors cette hypothèse est extra-théorique.

Que l'on considère ou non cette hypothèse comme une partie essentielle de la théorie, sa nature détermine ce que l'on a appelé avec Nagel le statut cognitif de la théorie,

¹⁰⁴Dire que la théorie a un contenu propositionnel ne revient pas à dire qu'elle est identifiable à des propositions linguistiques. C'est dire qu'elle prétend représenter le monde et que, en tant que telle, elle peut faire l'objet d'une attitude propositionnelle.

¹⁰⁵Voir note 101.

ou encore l'engagement épistémique de son utilisateur.¹⁰⁶ Cette hypothèse – intégrée ou non à la théorie – peut en effet prendre plusieurs formes. Un réaliste considérera que certains modèles décrits par les axiomes de la théorie représentent effectivement certaines portions du monde empirique ; un instrumentaliste se contentera d'affirmer qu'ils les représentent de telle façon qu'ils en permettent la prédiction, sans que cela signifie qu'à ces structures théoriques corresponde effectivement la structure sous-jacente du monde physique. Sur ce point, les défenseurs de la conception sémantique adoptent des positions différentes, que je n'examinerai pas ici.¹⁰⁷

Ce qui m'intéresse est plutôt la question du statut de cette hypothèse théorique : fait-elle, ou non, partie de la théorie ? Bas van Fraassen semble parfois suggérer que l'hypothèse théorique entre dans la caractérisation de la théorie, comme quand il reprend à son compte la formulation de Giere selon laquelle « une théorie consiste en (a) la *définition théorique*, qui définit une certaine classe de systèmes ; (b) une *hypothèse théorique*, qui affirme que certains (types de) systèmes réels appartiennent à cette classe » (van Fraassen, 1987, p. 111). Cependant, ni van Fraassen¹⁰⁸ ni Giere ne sont absolument fermes quant à l'intégration de l'hypothèse théorique dans la caractérisation de la théorie.

Or la question est loin d'être triviale : ne pas intégrer l'hypothèse théorique dans la caractérisation de la théorie, c'est faire de cette dernière un outil que l'on peut utiliser pour affirmer quelque chose sur le monde, mais qui en tant que tel n'affirme

¹⁰⁶ Plus précisément, si l'hypothèse est intégrée à la théorie, alors cela détermine son statut cognitif ; si on laisse l'hypothèse à l'extérieur, la théorie est inchangée et ce qui varie est l'engagement épistémique des agents.

¹⁰⁷ Giere (1988) et Suppe (1989) défendent des positions réalistes (un « réalisme constructif » pour le premier et un « quasi-réalisme » pour le second). Leurs arguments en faveur du réalisme ont été critiqués de façon assez convaincante, en particulier par Anjan Chakravartty (2001), qui montre que ces formes de réalisme sont plutôt des pseudo-réalistes, proches de la position instrumentaliste (« empirisme constructif ») défendue par van Fraassen. La position de van Fraassen (1980, p. 3) consiste à distinguer, au sein des modèles de la théorie, entre une « superstructure théorique » (cette expression est de Giere, 1988, p. 49, à propos de la distinction de van Fraassen) et des « sous-structures empiriques », seules ces dernières étant « candidates pour la représentation directe des phénomènes observables. Les structures qui peuvent être décrites dans les rapports expérimentaux et de mesure sont ce que l'on peut appeler les apparences : la théorie est empiriquement adéquate si elle a un modèle tel que toutes les apparences sont isomorphes aux sous-structures empiriques de ce modèle. » (van Fraassen, 1980, p. 64) Accepter une théorie, selon van Fraassen, consiste seulement à croire qu'elle est empiriquement adéquate en ce sens.

¹⁰⁸ Van Fraassen présente sa conception instrumentaliste tantôt comme une conception normative des théories, tantôt comme une description de l'attitude épistémique des scientifiques (voir van Fraassen, 1980, chap. 2 et 3). On peut mettre cette hésitation en relation avec celle concernant l'intégration ou non de l'hypothèse théorique dans la reconstruction formelle.

rien ; l'intégrer, c'est admettre que des aspects pragmatiques¹⁰⁹ de l'activité théorique entrent dans la caractérisation des théories elles-mêmes, et donc sortir d'une caractérisation strictement formelle. Ainsi Giere (1988) fait-il explicitement part de son hésitation à intégrer l'hypothèse théorique ; il craint que cela ne conduise à donner une image trop hétérogène des théories. En outre, cela reviendrait, selon lui, à trop insister sur les aspects linguistiques de l'utilisation des théories :

[On pourrait] dire qu'une théorie inclut à la fois les énoncés définissant la population de modèles et les hypothèses affirmant que plusieurs de ces modèles correspondent bien à certains types importants de systèmes réels. Le prix à payer pour avoir le beurre et l'argent du beurre est que la théorie devient une chose assez hétérogène. Elle inclut à la fois les définitions et les hypothèses empiriques. Mais c'est peut-être un prix assez bas quand il s'agit de saisir nos différentes intuitions de ce qu'est une théorie.

Ma seule objection à ce compromis est qu'il donne trop d'importance à des aspects linguistiques. Il concentre notre attention sur les énoncés qui définissent la population des modèles plutôt que sur les modèles eux-mêmes. Je préférerais substituer les modèles aux définitions. Les lois de Newton et les lois de force demeureraient, mais seulement de manière implicite, et non pas sous une forme linguistique. Elles seraient incarnées dans les modèles. (pp. 84-85)

Ce que Giere semble reconnaître ici en notant l'importance de ces « aspects linguistiques », c'est que l'analyse de la nature de cette hypothèse théorique passe par une attention à la manière dont les agents utilisent et comprennent, en pratique, les théories, et par conséquent à la forme particulière sous laquelle elles sont présentées. Plus encore qu'à donner une « image hétérogène » des théories, intégrer l'hypothèse théorique à la description du contenu des théories reviendrait ainsi à assigner des limites à sa reconstruction formelle.¹¹⁰

¹⁰⁹Une précision s'impose à propos du terme « pragmatique ». Comme on l'a vu (page 214), Hempel l'emploie pour désigner ce qui, dépendant du contexte et de l'utilisation des théories par les agents, ne fait pas partie de l'objet de la reconstruction formelle du philosophe. Van Fraassen (1980, chap. 5) développe une théorie pragmatique de l'explication scientifique, c'est-à-dire une théorie visant à montrer que ce qui compte comme une explication scientifique (indépendamment de la question de la psychologie de l'agent et du « sentiment de compréhension » que Hempel distingue de l'explication) est essentiellement dépendant du contexte. Mon emploi de ce terme n'est pas à interpréter en un sens technique : soit il me sert à désigner ce que Hempel ou van Fraassen entendent par là, quand je rapporte leurs propos, soit il sert simplement à désigner ce qui, dans l'activité théorique, relève de la pratique et de l'utilisation des théories.

¹¹⁰Notons que Giere est partisan d'une prise en compte des aspects pratiques et cognitifs de l'activité théorique, et qu'il cherche à défendre une version de la conception sémantique qui les intègre, en comprenant la notion de modèle à la fois en son sens logique et au sens que lui donnent les sciences cognitives dans l'étude du raisonnement. En ce sens, il ne nie pas l'importance du langage, et des représentations en général, dans l'activité théorique. Son rejet du « linguistique » est ici à comprendre,

Il semble en effet que la distinction entre structure mathématique et application empirique, en libérant l'analyse du contenu des théories des problèmes de philosophie du langage auxquels se sont trouvés confrontés les empiristes logiques, la dégage *ipso facto* de ceux relatifs à la valeur de vérité de ce contenu. L'analyse de ce dernier problème, en revanche, semble difficilement pouvoir échapper à celle de l'utilisation et du choix, par les agents, de telle ou telle formulation. Autrement dit, quand bien même la conception sémantique offre des outils permettant d'étudier le contenu des théories et leur application empirique en évitant la question de la référence linguistique, le problème du statut cognitif des théories peut difficilement être traité sans en passer par une attention à l'utilisation du langage – ou des représentations en général – qui établissent un lien entre elle et le monde.

Ce constat ne constitue pas, encore une fois, une objection au projet de la conception sémantique. Bas van Fraassen reconnaît explicitement le caractère pragmatique de l'acceptation d'une théorie (van Fraassen, 1980, p. 4) ; Patrick Suppes (1967, p. 66), quant à lui, appelle de ses vœux un développement des approches cognitives de l'utilisation du langage pour régler la question du réalisme et de l'instrumentalisme : ce n'est, selon lui, pas une question qui peut être traitée par une approche formelle – en tous cas pas tant que l'utilisation du langage ne fait pas l'objet de modèles scientifiques précis. Le but de la reconstruction formelle, pour lui, est de décrire les relations entre modèles, sans que cela implique un jugement sur la nature de l'engagement épistémique qui accompagne l'établissement de ces relations.¹¹¹ On aurait ainsi, d'une part, la reconstruction formelle du contenu des théories et de leur application empirique et, d'autre part, une analyse de la pratique scientifique et des attitudes épistémiques des agents.

En limitant ainsi le but qu'elle assigne à la reconstruction formelle, il semble que la conception sémantique soit à l'abri de toute critique visant à montrer qu'elle ne rend pas compte des aspects pratiques de l'utilisation des théories par les agents : de telles critiques n'atteignent pas leur cible, dans la mesure où elles cherchent à mettre la conception sémantique en défaut sur un terrain dans lequel elle ne prétend pas (toujours) s'aventurer. Mon but, à la section 3 du présent chapitre, est de montrer qu'il est cependant difficile, voire impossible, de distinguer clairement entre la caractérisation strictement formelle des théories et la manière dont elles sont utilisées par les agents. Autrement dit, ma critique ne consiste pas seulement à souligner le fait

selon moi, comme un rejet du modèle linguistique véhiculé par la conception syntaxique. Il n'est cependant pas clair que l'inclusion de l'hypothèse théorique dans la caractérisation de la théorie implique un retour à une conception « linguistique » au sens de l'empirisme logique.

¹¹¹Si on veut lui attribuer une position sur cette question, on pourra, comme Moulines et Sneed (1979, pp. 76-77), parler d'un « opérationnalisme raffiné ».

que la reconstruction formelle ne rend pas compte de ces aspects pratiques de l'utilisation des théories ; je vise à montrer, en outre, que l'ignorance même de ces aspects condamne l'entreprise de reconstruction formelle à un échec partiel, au regard de ses buts affichés. Auparavant, je me propose de préciser, dans la prochaine section, en quoi consistent ces « aspects pratiques » dont la conception sémantique ne peut pas (et, la plupart du temps, ne prétend pas) rendre compte.

2.4 À quoi sert la conception sémantique ?

La conception sémantique a fait l'objet, au cours des dernières décennies, de nombreuses critiques qui affirment qu'elle n'offre pas un traitement satisfaisant de plusieurs problèmes importants relatifs aux théories. Soazig Le Bihan (2006, 2008) examine les plus importantes de ces critiques et conclut qu'elles n'atteignent leur but que si l'on attend de la conception sémantique qu'elle résolve plus de problèmes qu'elle n'a jamais prétendu le faire. La conception sémantique est selon elle à l'abri de ces critiques si on la restreint à la prescription méthodologique d'outils pour étudier la structure des théories et si l'on n'y cherche ni une définition de la nature des théories¹¹² ni une analyse de la manière dont elles sont *utilisées pour représenter les phénomènes*. C'est ce dernier point que je souhaite examiner dans cette section.

Munie des outils formels de la théorie des modèles, la conception sémantique peut caractériser les relations structurelles (isomorphisme, plongement)¹¹³ entre modèles en vertu desquels certains peuvent être dits en représenter d'autres, mais le sens de la notion de représentation a ici le sens strictement formel qu'elle a dans la théorie des modèles. Elle ne désigne en aucun cas l'activité cognitive de l'agent qui utilise l'un de ces modèles pour en représenter un autre (ou représenter un modèle de données). On a vu (page 251) que les modèles, selon Suppes, ne représentent pas les phénomènes eux mêmes, mais que ces derniers sont eux-mêmes modélisés afin que l'on puisse

¹¹²Suppes affirme clairement que la conception sémantique ne vise pas à donner une définition de la nature des théories, qui permettrait d'énoncer des conditions nécessaires et suffisantes à l'affirmation selon laquelle un corps de connaissances est bien constitué en théorie, et d'en tracer nettement les contours : « Il ne me semble pas important de donner des définitions précises de la forme : X est une théorie scientifique si, et seulement si, ceci et cela. Ce qui importe est de reconnaître que l'existence d'une hiérarchie de théories s'élevant depuis la méthodologie de l'expérimentation vers le test de la théorie fondamentale est un ingrédient essentiel de toute discipline scientifique sophistiquée. » (Suppes, 1967, pp. 63-64) Le but de la conception sémantique n'est donc pas de forcer les théories à entrer dans un « moule » prédéfini. Elle doit au contraire fournir des outils permettant d'explorer le contenu de types différents de corps de connaissance dont tout le monde s'accorde à dire que ce sont des théories. Sur ce point, voir l'analyse de Frederick Suppe (1977a, pp. 62-66), et la taxonomie des théories proposée par Anatol Rapoport (1958).

¹¹³Voir page 247.

établir une relation de représentation entre les modèles théoriques et les modèles de données. Or, l'activité qui consiste à modéliser les phénomènes et à proposer, comme le dit Suppes (1960, p. 300), une « version hautement schématisée de l'expérience », c'est-à-dire à représenter, par exemple, le balancier d'une horloge normande *comme* un pendule simple, ne peut pas être décrite au moyen d'outils strictement formels. Autrement dit, la description de l'activité théorique au moyen de ces seuls outils ne permet pas de rendre compte de l'aspect essentiellement *intentionnel* de la représentation comme activité cognitive (voir Frigg, 2002; Suárez, 1999, 2003). Elle ne permet de décrire ni comment les phénomènes sont structurés en modèles de données, ni comment les agents utilisent ces modèles et ces relations entre modèles pour tirer des inférences à propos des phénomènes qu'ils étudient.

Les différents critiques de la conception sémantique s'accordent à dire que, pour traiter de ces questions propres à l'aspect propositionnel des théories, ainsi que de celle, évoquée à la section précédente, de l'engagement réaliste qui accompagne ou non l'utilisation des théories pour représenter les phénomènes, le philosophe doit prêter attention à la nature de la représentation et aux questions linguistiques de dénotation et de référence qui l'accompagnent. Les défenseurs de la conception sémantique soulignent, comme on l'a vu (voir page 251), que la notion de règle de correspondance est une simplification abusive de la manière dont les théories sont appliquées aux phénomènes. Cependant, tout en proposant des outils qui permettent de mieux rendre compte de la complexité de la relation entre structure théorique et phénomènes, ils offrent aussi, selon leurs critiques, une image de l'activité théorique qui reste à un très haut niveau d'abstraction, et qui ne permet en aucun cas de comprendre comment, en pratique, les théories ou les modèles sont utilisés pour représenter et, par là, expliquer et prédire les phénomènes. L'adoption du postulat, qui est au fondement de l'approche sémantique, selon lequel les modèles utilisés par les scientifiques peuvent être définis comme des modèles logiques (voir page 249), revient ainsi à laisser de côté une partie importante du fonctionnement de ces modèles.¹¹⁴

Il est à peu près clair que Suppes et van Fraassen¹¹⁵ ne prétendent pas rendre compte de cet aspect de l'activité théorique au delà de ce que les ressources formelles permettent.¹¹⁶ Aussi, cette critique de la conception sémantique est essentiellement

¹¹⁴Pour les critiques de la conception sémantique fondées sur le refus de ce postulat, voir, entre autres, (Suárez, 1999, 2003, 2004; Frigg, 2002, 2006; Thomson-Jones, 2006).

¹¹⁵Il convient de noter que van Fraassen (2008) vient tout juste de publier un ouvrage où il aborde ces problèmes de front, et étudie le détail de l'utilisation des représentations dans la pratique scientifique. Il n'affirme aucunement qu'une telle étude doive s'insérer dans le projet de reconstruction formelle.

¹¹⁶Giere, en revanche, en plaçant sa compréhension de la conception sémantique dans une pers-

externe, et ne suffit pas à en montrer les limitations *internes*. Je reviendrai, dans la suite de ce travail (chapitre 6, section 2.2), sur les critiques qui ont été adressées, sur ce point, à la conception sémantique. Je montrerai en particulier qu'elles ne sont pas allées elles-mêmes assez loin dans l'élaboration des outils nécessaires pour rendre compte de l'interaction cognitive des scientifiques avec les représentations concrètes qu'ils utilisent, et sont restées trop attachées à une définition abstraite de la notion de modèle et de représentation.

Pour l'heure, la question qui m'intéresse est de savoir si et dans quelle mesure, en restreignant comme le suggère Soazig Le Bihan le but de la conception sémantique, et donc en ignorant délibérément la manière dont les agents utilisent les modèles, on peut affirmer qu'elle est couronnée de succès. Si, sur ce point, on rencontre des difficultés, alors on aura établi une critique authentiquement *interne* de la conception sémantique. C'est cette question que je souhaite examiner à présent ; au terme de cet examen, on verra que l'entreprise de reconstruction formelle du contenu des théories *ne peut pas ignorer*, comme la conception sémantique le prétend, ces aspects pratiques et cognitifs.

3 Formulations et reconstruction formelle : l'identité de la mécanique classique selon les approches formelles

Dans cette section, je souhaite montrer les limites internes du projet de la conception sémantique et des approches formelles en général. Le but d'un tel argument est de montrer que le projet global qui anime mon travail dans son ensemble ne consiste pas uniquement en l'adoption d'une perspective différente de celle des approches formelles, mais trouve une justification dans ces limites internes.

Les difficultés rencontrées par le projet de reconstruction syntaxique des empiristes logiques ont été mises en évidence à la fin de la section 1, et elles ont motivé la démarche adoptée par la conception sémantique, présentée à la section 2. Afin d'évaluer le succès de cette dernière, je propose dans un premier temps (section 3.1) de revenir sur les buts principaux de la reconstruction axiomatique des théories, tels que Patrick Suppes (1968) les envisage. Ces buts, et les présupposés mêmes qui en sous-tendent la poursuite, sont pour la plupart communs aux deux approches formelles. On verra pourtant, dans un deuxième temps (section 3.2), que les défenseurs

pective cognitive, prétend explicitement rendre compte, dans ce cadre, de ces aspects cognitifs et pragmatiques de l'utilisation des modèles comme représentations. C'est même un des avantages, selon lui, de la conception sémantique sur l'approche syntaxique fondationnelle : « ces études fondationnelles n'ont eu presque aucun impact sur la manière dont la science est enseignée, étudiée, ou comprise par les scientifiques. » (Giere, 1988, p. 63)

de la conception sémantique reprochent souvent au projet formaliste de la conception syntaxique de ne pas permettre de distinguer entre le contenu d'une théorie et sa formulation ; je montrerai que cette critique repose sur une confusion entre formulation (dans tel ou tel langage mathématique) et formalisation (au sens strict). On verra finalement (section 3.3), par l'examen de la manière dont la conception sémantique permet de traiter le problème de la diversité des formulations de la mécanique classique, que cette confusion entre formulation et formalisation est symptomatique d'une de ses limites internes.

3.1 Les vertus de la reconstruction formelle

La conception sémantique se propose d'utiliser les ressources formelles de la théorie des modèles pour étudier le contenu des théories scientifiques ; une axiomatisation rigoureuse doit permettre au philosophe de procéder à la clarification conceptuelle des théories par-delà les aspects inessentiels de leur(s) formulation(s) habituelle(s). Cette clarification conceptuelle, qui doit aider au traitement de problèmes philosophiques aussi bien que scientifiques est, pour Suppes comme pour Carnap, la tâche primordiale du philosophe des sciences :

Le rôle de la philosophie des sciences est de clarifier des problèmes conceptuels et d'expliciter les présupposés fondamentaux de chaque discipline scientifique. La clarification de problèmes conceptuels ou la construction d'un fondement logique explicite sont des tâches qui ne sont de nature ni tout à fait empirique, ni mathématique. On peut les considérer comme des tâches proprement philosophiques qui sont directement pertinentes pour la science (Suppes, 1968, p. 653).

Conceptions syntaxique et sémantique s'accordent à considérer que cette entreprise passe par l'élimination du « bruit » lié à certains aspects des formulations habituelles des théories. La psychologie du chercheur, les types de représentations favorisés et les présupposés implicites d'une communauté scientifique sont autant d'aspects que la reconstruction formelle cherche à contourner pour explorer le contenu des théories.

Dans un article intitulé « The desirability of formalization in science », Suppes (1968) énumère les différentes vertus de l'axiomatisation¹¹⁷ des théories. Elles correspondent aux buts poursuivis par toute entreprise formelle ; l'axiomatisation dans

¹¹⁷Suppes, qui parle généralement d'« axiomatisation » des théories (par exemple, Suppes, 1957), parle ici de « formalisation ». C'est pourtant bien de la formalisation au sens large qu'il s'agit, incluant les reconstructions formelles proposées par la conception sémantique (voir page 192). Je me suis efforcée, jusqu'ici, de n'employer le terme « formalisation » que quand il s'agissait de la formalisation stricte des empiristes logiques, et de parler de « reconstruction formelle » dans les autres cas. Par commodité, je suivrai dans cette section l'usage de Suppes, en tâchant de dissiper les ambiguïtés possibles.

les termes de la théorie des ensembles est cependant selon lui la mieux à même de remplir sa mission, pour les raisons présentées dans ce qui précède. Selon Suppes, formaliser (i.e. reconstruire formellement) une théorie est le moyen de rendre parfaitement explicites les concepts utilisés dans les théories (Suppes, 1968, p. 654)¹¹⁸, et permet d'envisager une standardisation du langage de la science facilitant communication entre disciplines et apprentissage scientifique et ouvrant ainsi une voie réelle vers l'unité de la science (Suppes, 1968, p. 654).¹¹⁹ De plus, isoler l'ensemble des hypothèses indépendantes [*self-contained assumptions*] d'une théorie permet d'éviter les verbalisations *ad hoc* (Suppes, 1968, p. 655). Enfin et surtout, Suppes affirme que, en isolant les hypothèses minimales indispensables à l'expression d'une théorie, une bonne axiomatisation permet d'atteindre un niveau de généralité et d'objectivité autrement inaccessible. Ses propos ne sont pas sans rappeler le projet hertzien de présentation des principes de la mécanique examiné au chapitre 3 :

La formalisation d'une théorie rend possible l'analyse objective des hypothèses minimales indispensables à l'énoncé de la théorie. Il y a, je pense, un attrait directement esthétique dans le fait de trouver l'ensemble des hypothèses mutuellement indépendantes qui sont autonomes dans la formulation d'une théorie. [...]
L'aspect significatif de la question : « quelles sont les hypothèses minimales dans les termes desquelles la théorie peut être formulée ? » est un test assez direct du sérieux de notre compréhension de la théorie, et une mesure de la profondeur de son développement. (Suppes, 1968, p. 656)

Plutôt que l'établissement d'un critère de scientificité auquel aucune théorie, de fait, ne satisfait pleinement, l'axiomatisation selon la méthode sémantique cherche à présenter de la manière la plus explicite et la plus simple possible le contenu des théories existantes. Pour Suppes, tout comme pour Hertz¹²⁰, ce sont nos connaissances scientifiques en leur état actuel qu'il s'agit de présenter d'une manière qui nous en donne la compréhension la plus globale, en *réduisant les effets de perspective* liés à la manière dont on les formule et dont on les aborde habituellement. Une bonne axiomatisation nous « offre souvent le moyen de voir la forêt au lieu des arbres », éliminant « les traits provinciaux et inessentiels de la manière dont une théorie a été pensée » (Suppes, 1968, p. 654). Axiomatiser une théorie, c'est donc se donner le moyen d'avoir une compréhension claire et directe de ses hypothèses et de ses

¹¹⁸L'exemple pris est celui du concept de probabilité, confus jusqu'à son axiomatisation par Kolmogorov en 1933.

¹¹⁹Il s'agit ici d'un avantage pour les scientifiques plus encore que pour les philosophes. L'axiomatisation d'une théorie semble donc pouvoir prétendre à devenir sa formulation standard.

¹²⁰Le projet de Hertz dans les *Principes* est en effet d'« exposer un agencement parfaitement déterminé des lois de la mécanique qui soit compatible avec l'état actuel de notre savoir [...] ». (Hertz, 1894, p. 63). Voir chapitre 3, section 4.2.

concepts fondamentaux. Notons que l'axiomatisation ainsi conçue fournit aussi un test de l'état d'avancement d'une science : « exiger qu'une branche de la physique soit formulée dans les termes d'un ensemble minimal d'hypothèses ou d'axiomes est une requête sensée », tandis que « la question de ce qui constitue l'ensemble d'hypothèses minimal mais suffisant de la psychanalyse ne peut pas recevoir, à l'heure actuelle, de réponse sérieuse » (p. 656).

Insistons sur le point suivant : s'il s'agit bien, pour Suppes, de dégager le contenu des théories existantes et non pas de dessiner les contours de ce que devrait être une théorie pour être parfaitement scientifique au moyen d'une formalisation stricte à la manière des empiristes logiques, il ne s'agit pas non plus de se contenter des formulations habituelles, sous lesquelles les théories *semblent dire* des choses qu'elles ne disent pas vraiment. Suppes affirme ainsi avoir montré, dans (Suppes, 1959), que, contrairement aux apparences, la cinématique de la théorie de la relativité restreinte repose sur une seule hypothèse fondamentale.¹²¹ Ce faisant, il a montré que, derrière le manque de clarté des présentations habituelles de la théorie de la relativité restreinte, ce qu'elle dit, et ce à quoi un scientifique qui l'utilise est minimalement engagé, se réduit à une seule hypothèse, malgré ce qu'en pensent habituellement les utilisateurs de cette théorie :

Une réponse possible [au constat du manque de clarté des hypothèses fondamentales de la théorie de la relativité restreinte] est que [cette théorie] est beaucoup de choses pour beaucoup de physiciens, et que l'on ne devrait pas attendre des philosophes qu'ils donnent un ensemble clair et fondamental d'hypothèses desquelles tous les résultats souhaités découlent. Pour ma part, je ne pense pas que cette opinion soit correcte. (Suppes, 1968, p. 657)

S'il ne s'agit pas de forcer la théorie à rentrer dans un moule auquel elle n'est pas adaptée, c'est bien à dégager ce qu'elle dit vraiment, derrière ce qu'en comprennent les utilisateurs et ce qu'en suggèrent les formulations habituelles, que la reconstruction formelle vise.

Dégager ainsi le contenu et les hypothèses minimales d'une théorie présente un avantage majeur dans les cas où il existe plusieurs formulations standard d'une théorie ; l'axiomatisation permet en effet de prouver que deux formulations qui semblent dire des choses différentes décrivent en fait les mêmes structures mathématiques. Ainsi, les formulations ondulatoire et matricielle de la mécanique quantique sont

¹²¹Cette hypothèse est celle selon laquelle « tous les observateurs au repos dans des référentiels inertiels obtiennent les mêmes mesures des distances relativistes le long des trajectoires inertielles quand leurs instruments de mesure sont calibrés de façon identique. Le fait que les observateurs dans des référentiels différents se meuvent à une vitesse uniforme les uns par rapport aux autres est une conséquence et *non* une hypothèse de cette analyse. » (Suppes, 1968, p. 657)

identiques, au sens où « il y a un unique espace hilbertien complexe, séparable, à un isomorphisme près, et toutes deux sont des réalisations de cet espace » (Suppes, 1968, p. 655). C'est en ce sens que l'axiomatisation permet de dégager le contenu *objectif* (Suppes, 1968, p. 655) d'une théorie : derrière la diversité des formulations, qui semblent parfois s'opposer sur des principes fondamentaux, on peut dégager une invariance structurelle dans les modèles expérimentaux que l'on peut en tirer.

Ces différents avantages de l'axiomatisation sémantique, que ce soit dans les termes de la théorie des ensembles ou dans ceux de la description de trajectoires dans un espace des phases¹²², ont pu être testés par la mise en œuvre effective d'axiomatisations de théories existantes. En tant que programme d'axiomatisation des théories, la conception sémantique a indéniablement fait ses preuves. J'ai mentionné (page 247) la possibilité d'étudier la question des relations interthéoriques de manière fructueuse à l'aide de l'établissement de théorèmes de représentation ; Suppes (2002) applique cette méthode à de nombreux domaines scientifiques (théorie des probabilités, théories de l'espace-temps, mécanique classique et quantique, théories du langage). Dans le domaine de la physique, et plus précisément de la mécanique quantique, les travaux de Hughes (1989) ainsi que ceux de van Fraassen (1991), par exemple, ont montré que la conception sémantique permet d'éclairer certains problèmes scientifiques internes à cette théorie. En biologie, plusieurs axiomatisations de la théorie de l'évolution et de la génétique des populations ont été proposées (voir par exemple Beatty, 1982; Lloyd, 1988; Thompson, 1989, 2007), permettant de clarifier certains concepts et d'examiner le statut de différents principes, dans un domaine où la notion de loi est considérée comme hautement problématique et comme ne permettant pas de rendre compte du type d'explication et de confirmation à l'œuvre. En tant que programme de recherche, la conception sémantique prouve ainsi par la pratique une de ses supériorités proclamées sur la conception empiriste logique de la science : elle fournit des outils d'analyse efficaces des théories scientifiques existantes plutôt que de décrire la forme idéale que devrait revêtir toute théorie digne de ce nom – mais qu'aucune ne revêt de fait.

Dans ce qui suit, je propose de nuancer l'optimisme des défenseurs de la conception sémantique : malgré ces succès indéniables, je vais montrer qu'il n'est pas évident que l'axiomatisation d'une théorie suffise à distinguer ce que la théorie dit vraiment de ce que ses formulations habituelles semblent lui faire dire ; je vais montrer, en reprenant l'exemple de la mécanique classique, que la méthode axiomatique de la conception sémantique ne suffit pas à déterminer de manière objective et indépendante de la

¹²²Thompson (2007) montre que la génétique mendélienne peut être axiomatisée sous les deux formes.

perspective particulière des utilisateurs si deux formulations différentes d'une théorie disent bien la même chose. Auparavant, je souhaite souligner une ambiguïté, dans les propos de certains défenseurs de la conception sémantique, à propos de ce qu'il faut entendre par « formulation » d'une théorie.

3.2 Formaliser pour aller « au delà » des formulations : conception syntaxique et conception sémantique

Comme je viens de le rappeler, un des avantages notables que ses défenseurs attribuent à la conception sémantique est d'offrir les moyens de dépasser les différences inessentiels et « provinciales » liées aux formulations des théories et de dégager leur contenu objectif, permettant ainsi de montrer, entre autres, que des théories qui « disent » apparemment des choses différentes sont en fait des descriptions des mêmes structures mathématiques. Alors qu'il semble que ce but – sinon sa réalisation – soit commun aux deux approches formelles, les défenseurs de la conception sémantique¹²³ accusent souvent la conception syntaxique d'une confusion grave entre théorie et formulation de théorie, et affirment que la formalisation qu'elle propose ne permet pas de distinguer entre les deux. Je vais montrer que cette accusation est infondée, et qu'elle repose, chez ceux qui la portent, sur une ambiguïté dans l'usage même de la notion de « formulation ».

Selon Suppe, une des supériorités de la conception sémantique sur la conception syntaxique est qu'elle permet de montrer qu'une théorie peut recevoir plusieurs formulations différentes, sans que cela affecte son identité :

Selon la conception sémantique des théories, les théories scientifiques ne sont pas des entités linguistiques, mais plutôt des entités ensemblistes [...]. Cela contraste nettement avec plusieurs versions de la conception orthodoxe des théories, qui conçoit les théories comme des systèmes axiomatiques partiellement interprétés, et donc comme des entités linguistiques. Dire que quelque chose est une entité linguistique implique que des changements dans ses caractéristiques linguistiques, y compris la formulation de son système d'axiomes, produit une nouvelle entité. Par conséquent, selon la conception orthodoxe, un changement dans la formulation d'une théorie est un changement de théorie. (Suppe, 1989, pp. 3-4)

Un de ses exemples favoris, à ce propos, est d'ailleurs celui de la mécanique classique : alors que la conception sémantique permet, selon Frederick Suppe, de rendre compte de l'identité de cette théorie derrière la diversité de ses formulations, la conception syntaxique est condamnée à affirmer qu'il y a autant de théories différentes que

¹²³Il s'agit principalement de Frederick Suppe et de Bas van Fraassen. À ma connaissance, Patrick Suppes ne porte jamais ce genre d'accusation contre la conception syntaxique.

de formulations, en raison du fait qu'elle définit les théories comme des « entités linguistiques » :

Les théories, telles qu'elles sont actuellement utilisées par les chercheurs, admettent un certain nombre de formulations linguistiques différentes – par exemple, la mécanique classique du point reçoit parfois une formulation lagrangienne et parfois une formulation hamiltonienne – mais c'est la même théorie, quelle que soit la formulation utilisée. Pour cette raison, les théories ne peuvent pas être identifiées à leurs formulations linguistiques ; ce sont plutôt des entités extralinguistiques auxquelles il est fait référence et qui sont décrites par leurs diverses formulations linguistiques. Ces structures sont des *modèles métamathématiques* de leurs formulations linguistiques, et la même structure peut être le modèle de plusieurs ensembles différents d'énoncés ou formulations linguistiques de la théorie [...]. (Suppe, 1989, p. 82)

Bas van Fraassen tient par endroits le même type de propos que Frederick Suppe : « L'erreur, je pense, a été de confondre une théorie avec la formulation d'une théorie dans un langage particulier. » (van Fraassen, 1987, p. 109) Il voit dans l'identification d'une théorie à un énoncé le germe d'une tension interne à la conception syntaxique :

[La conception traditionnelle des théories] affirme qu'une théorie est juste un énoncé très complexe. Deux énoncés peuvent être logiquement équivalents, auquel cas ils « disent la même chose ». Par conséquent deux axiomatisations d'une théorie, étant logiquement équivalentes, peuvent être appelées deux énoncés ou formulations de la même théorie. Une telle conception traditionnelle comporte des tensions. La locution « deux énoncés peuvent être des formulations de la même théorie » elle-même suggère qu'une théorie n'est pas du tout un énoncé, mais seulement quelque chose qui est formulé ou exprimé au moyen d'un énoncé. (van Fraassen, 1989, p. 190).

Ce que dénoncent ici Suppe et van Fraassen, c'est l'idée selon laquelle une théorie *est* un ensemble d'énoncés, dont la moindre modification syntaxique implique une modification essentielle ; une telle thèse conduirait non seulement à soutenir que différentes formulations mathématiques sont différentes théories, mais encore l'idée bien étrange selon laquelle les formulations de la mécanique en anglais et en français, par exemple, n'expriment pas la même théorie. Cette critique repose cependant, selon moi, sur une erreur quant au but et à la nature même de la formalisation des théories envisagée par les empiristes logiques.

Carnap, on l'a vu, cherche à dégager des critères de scientificité et de signification cognitive pour les théories, ce qui le conduit à l'idée d'une forme canonique sous laquelle toute théorie devrait idéalement pouvoir être reconstruite. Ce faisant, son but est d'identifier et de garantir, en le reconstruisant de façon appropriée, le contenu cognitif de la théorie, et non d'affirmer que les énoncés au moyen desquels

on le reconstruit *sont* la théorie.¹²⁴ L'affirmation (vraie) selon laquelle la conception syntaxique décrit les théories comme des entités linguistiques signifie simplement qu'elle en reconstruit le contenu au moyen d'énoncés non interprétés directement reliés au monde empirique, qui acquièrent une signification de cette relation même. Comme je l'ai souligné à la section 1, ces formes d'énoncés, avant toute interprétation, *ne ressemblent pas du tout à des énoncés du langage naturel, même étendu au langage scientifique*. Aussi, l'accusation que Frederick Suppe porte à la conception syntaxique, selon laquelle elle opère une identification entre théorie et formulation de théorie semble reposer sur une confusion entre langage naturel (étendu au langage scientifique) et langage formel, ou plus précisément entre la formulation d'une théorie et sa formalisation logique.

Les empiristes logiques n'affirment pas plus que les défenseurs de la conception sémantique que les deux formulations de la mécanique, par exemple, expriment des théories différentes : elles peuvent très bien être des présentations différentes, en langage naturel, du même squelette syntaxique, ou de deux squelettes syntaxiques dont les règles de la logique doivent permettre de montrer l'équivalence. La « tension » que van Fraassen propose de voir dans la position des empiristes logiques n'en est pas une : de même que deux énoncés logiquement équivalents expriment la même

¹²⁴ Affirmer que la conception syntaxique a fait une « erreur de catégorie » (Le Bihan, 2006) en identifiant une théorie à un ensemble d'énoncés (une entité linguistique) et que la conception sémantique a raison de l'identifier à un ensemble de modèles (une entité extra-linguistique) relève à mon avis d'une confusion sur le but de la conception sémantique autant que sur celui de la conception syntaxique. Le but de la reconstruction formelle n'est en effet pas d'émettre une thèse sur le *type d'objets* que sont les théories scientifiques ou, pour le dire autrement, sur la classe ontologique à laquelle elles appartiennent, permettant d'isoler des objets et de leur appliquer l'étiquette « théorie » en fonction de certains critères. Voir à ce sujet l'introduction de l'article de Patrick Suppes (1967) intitulé « What is a scientific theory ? » : « Souvent, quand on demande ce qu'est telle ou telle chose, on attend une réponse claire et définie. Si, par exemple, quelqu'un me demande ce qu'est un nombre rationnel, je peux donner la réponse simple et précise selon laquelle un nombre rationnel est le ratio de deux entiers. Il existe d'autres types de questions simples pour lesquelles une réponse précise peut être donnée mais pour lesquelles dans le discours ordinaire une réponse plutôt vague est habituellement donnée et acceptée. Quelqu'un lit quelque chose à propos des nectarines dans un livre, mais n'a jamais vu de nectarine, ou a peut-être vu des nectarines mais ne connaît pas leur nom français [anglais]. Il pourra me demander : "Qu'est-ce qu'une nectarine ?" et je répondrai probablement : "Une sorte de pêche à la peau lisse." Certes, ce n'est pas une réponse très exacte, mais si la personne qui me pose la question sait ce que sont les pêches, elle pourrait s'en satisfaire. Le type de question que je veux discuter ici n'entre dans aucun de ces deux schémas. Les théories scientifiques ne sont pas comme les nombres rationnels ou les nectarines. Elles ne sont évidemment pas comme les nectarines, parce que ce ne sont pas des objets physiques simples. Elles sont comme les nombres rationnels au sens où elles ne sont pas des objets physiques, mais elles sont très différentes des nombres rationnels au sens où elles ne peuvent pas être définies d'une manière simple et directe comme des objets abstraits, non physiques. » (Suppes, 1967, p. 55)

proposition, deux ensembles d'énoncés peuvent exprimer la même théorie.¹²⁵

Si la conception sémantique peut adresser une critique justifiée à la conception syntaxique, ce devrait plutôt être l'inverse de celle mentionnée ci-dessus. Comme on l'a vu à la section 1, les empiristes logiques, en voulant distinguer le squelette syntaxique d'une théorie de son interprétation, rendent impossible la reconstruction de ce contenu. C'est parce qu'ils négligent le fait, montré à la section 1.6.4, qu'une théorie est toujours déjà interprétée, *parce qu'elle est toujours déjà formulée dans un certain langage*, que les empiristes logiques sont condamnés à faire d'une théorie scientifique un ensemble de formules abstraites qui ne disent ni ne peuvent rien dire. Loin d'identifier théorie et formulation, la conception syntaxique prononce un véritable divorce entre les théories et leurs formulations linguistiques, qui rend impossible d'en retrouver le contenu. Autrement dit, le « péché » de la conception syntaxique, loin de résider dans une identification entre théorie et formulation linguistique, consiste dans l'extrême inverse, c'est-à-dire dans le refus de reconnaître, comme Nagel, et comme les défenseurs de la conception sémantique, qu'une théorie est toujours déjà interprétée. Elle est donc totalement insensible aux différences de formulation, contrairement à ce que les propos de Frederick Suppe et de Bas van Fraassen rapportés ci-dessus laissent entendre : non seulement elle permet de dire que les différentes formulations de la mécanique sont équivalentes, mais encore elle ne « voit » même pas leurs différences.

Par conséquent, le traitement des formulations linguistiques comme inessentiels au contenu des théories non seulement ne différencie pas les deux conceptions étudiées mais c'est même un de leurs points communs les plus fondamentaux, qui leur vaut le nom d'approches « formelles ». Cependant, comme je l'ai souligné, les outils qu'elles mettent en œuvre ne sont pas les mêmes. À la formalisation stricte, la conception sémantique substitue une reconstruction axiomatique qui se donne toutes les notions mathématiques nécessaires. Or, ce que je voudrais montrer dans la dernière section de ce chapitre, c'est que, dès lors que l'on renonce à la formalisation stricte (dont on a vu les limites), il n'est plus vraiment possible de distinguer entre la reconstruction formelle d'une théorie et sa formulation. Autrement dit, les axiomatisations des théories proposées par la conception sémantique peuvent être elles-mêmes considérées comme des formulations de ces théories – sans doute plus cohérentes, plus claires et plus explicites que leurs formulations habituelles, mais pas d'une nature essentiellement différente.¹²⁶

¹²⁵Voir (Quine, 1975) pour une analyse de la notion d'équivalence appliquée au cas des différentes formulations d'une même théorie. Le critère qu'il vise à élaborer devrait assurément permettre de traiter les formulations de la mécanique comme des formulations équivalentes.

¹²⁶Ce constat permet d'interpréter la confusion de Suppe, mentionnée ci-dessus, entre formulation et formalisation, comme une conséquence de l'absence de distinction nette entre la formalisation

3.3 La reconstruction formelle suffit-elle à établir l'identité des formulations de la mécanique classique ?

On a vu, à la section 1.6.5, que l'ambiguïté de la position de Nagel à propos de la mécanique est symptomatique d'un problème interne à sa conception des théories, lui-même significatif des limites de la reconstruction des théories par les empiristes logiques. Que nous en dit la conception sémantique, quand elle cherche à axiomatiser la mécanique, et non pas simplement à invoquer cet exemple contre la conception syntaxique ? Offre-t-elle véritablement les outils d'une reconstruction formelle de son contenu qui mette en évidence l'équivalence de ses différentes formulations ?

Avant de répondre à cette question, il convient de dire un mot du statut de l'exemple de la mécanique dans la présente argumentation. Les défenseurs de la conception sémantique, on l'a vu, prennent souvent la diversité des formulations de la mécanique comme l'exemple d'un type de problème que leur méthode axiomatique permet de résoudre, en montrant l'équivalence de ces formulations. S'il s'avère que ce n'est pas le cas, cela ne suffit pas pour autant à condamner le projet axiomatique : il se pourrait que les différentes formulations de la mécanique, contrairement à ce que l'on pense, ne soient pas équivalentes. Il se pourrait aussi que les hypothèses et les concepts fondamentaux de cette théorie soient moins bien définis qu'on ne croit habituellement, et offrent une résistance à l'entreprise axiomatique. Autrement dit, contre ce que suggère l'idée persistante jusqu'à la fin du dix-neuvième siècle selon laquelle la mécanique classique serait le standard même de ce que l'on appelle une théorie scientifique, on peut affirmer que l'échec éventuel de son axiomatisation n'est pas une preuve de l'échec du projet axiomatique en général.

Toutefois, on est en droit d'attendre une réponse claire de la part de la conception sémantique. D'après les vertus que ses défenseurs lui attribuent, elle doit être en mesure de discriminer entre une différence de formulation, sans conséquence sur le contenu de la théorie, et une différence de contenu, que la reconstruction formelle doit mettre en évidence ; elle doit montrer soit que les différentes formulations de la mécanique décrivent des structures isomorphes, soit que seulement certaines sous-structures sont isomorphes, soit enfin qu'il est pour l'heure impossible de dégager un ensemble d'hypothèses fondamentales la caractérisant, tout comme dans le cas de la psychanalyse (voir page 262). Cette dernière option n'est, on s'en doute, choisie par personne. La mécanique classique est une des premières théories à avoir fait l'objet d'axiomatisations suivant la méthode sémantique (voir McKinsey *et al.*, 1953; Suppes,

d'une théorie au sens strict et sa reconstruction formelle. Cette dernière, opérée dans un langage qui comporte déjà une interprétation mathématique, ne peut pas être aussi clairement distinguée des formulations de la théorie.

1957). On va cependant voir que la conception sémantique a une position au moins aussi ambiguë que celle de Nagel (1961) sur le statut des différentes formulations de la mécanique classique, et que cette ambiguïté est symptomatique des limites internes du projet axiomatique, en tant qu'il prétend établir une distinction nette entre ce qui relève du contenu d'une théorie et ce qui relève de sa formulation.

Prenons pour commencer l'axiomatisation de la mécanique du point proposée par Suppes (1957), partiellement reproduite ci-dessus (page 243). Il s'agit clairement d'une axiomatisation *de la formulation newtonienne* de la mécanique. En effet, les variables correspondent aux concepts de force et de masse, et il est fait usage de la notion de dérivée seconde par rapport au temps dans l'expression des axiomes fondamentaux. Dans la mesure où le langage dans lequel l'axiomatisation est faite est inessentiel – c'est un des slogans de la conception sémantique –, on choisit les notions primitives et les outils mathématiques qui sont les plus commodes pour le but visé. Pour reconstruire le contenu de la mécanique newtonienne, il est naturel de se donner les outils mathématiques qu'elle utilise ; c'est précisément une des supériorités de la conception sémantique sur celle des empiristes logiques, on l'a vu, que de ne pas épurer le langage des théories au point de n'admettre qu'une syntaxe non interprétée et des outils logiques élémentaires. Suppes ne prétend pas isoler et individuer les théories en en donnant une et une seule reconstruction formelle, mais bien examiner les relations entre les structures décrites par différentes théories, ou par différentes formulations d'une théorie. Si les différentes formulations de la mécanique sont équivalentes, la méthode axiomatique doit permettre de le montrer, non pas en proposant une et une seule axiomatisation de *la* mécanique (indépendante de chacune de ses formulations), mais en permettant de comparer les structures mathématiques décrites par chacune de ses formulations – et donc en proposant une axiomatisation de chacune de ces formulations. Autrement dit, l'axiomatisation d'une théorie ne se distingue pas fondamentalement d'une reformulation permettant d'en clarifier le contenu mathématique. La question que je pose ici est la suivante : l'axiomatisation des différentes formulations de la mécanique permet-elle de montrer l'équivalence (ou d'identifier clairement la différence) de leur contenu ?

Balzer, Moulines et Sneed (1987), adeptes d'une forme d'axiomatisation inspirée de la méthode ensembliste de Suppes¹²⁷ ont cherché à répondre à cette question. Pour cela, ils proposent une axiomatisation de la mécanique newtonienne (pp. 103-108) d'une part et de la mécanique lagrangienne (pp. 149-155) de l'autre. Ils affirment considérer « la mécanique lagrangienne comme une théorie distincte à ne pas identifier à la mécanique classique. Ses concepts de base et sa loi fondamentale sont différents

¹²⁷Pour un exposé récent de cette méthode, voir (Moulines, 2002, 2006).

de ceux de [la mécanique newtonienne du point] » (p. 149). Ils montrent ensuite que, si leur équivalence empirique peut être démontrée par la spécification de leur domaine d'application, qui trivialisait en fait cette équivalence (pp. 292-295), leur équivalence complète est « très loin d'être triviale » (p. 303). Les trois tentatives de démonstration qu'ils proposent échouent, et ils concluent qu'on ne peut affirmer leur équivalence.

Cela appelle plusieurs commentaires. En premier lieu, leur affirmation initiale selon laquelle mécaniques newtonienne et lagrangienne sont deux théories différentes, dont on peut essayer de montrer l'équivalence, révèle un usage particulier de la notion de théorie, selon lequel deux théories, même strictement équivalentes, restent deux théories distinctes si leurs lois et concepts fondamentaux sont différents. Dans le cas où les deux théories s'avéreraient équivalentes (ce que Balzer, Moulines et Sneed n'excluent pas d'emblée), ces différences ne concerneraient que leurs formulations particulières. Les prendre en compte dans l'identification d'une théorie, comme ils le font, c'est donc prendre en compte des aspects que les approches formelles visent en principe à évacuer. Accorder de l'importance au fait que la formulation newtonienne parle de forces quand la formulation lagrangienne parle d'énergie, alors même que les structures décrites sont peut-être identiques, c'est considérer que le langage dans lequel une théorie est formulée et la manière dont le sujet qui l'utilise la relie au monde empirique a une importance dans l'identification d'une théorie. Cette remarque, pour l'heure, n'a pas d'autre valeur que de pointer, chez ces auteurs, la reconnaissance des limites de toute reconstruction formelle.¹²⁸ Elle prendra bientôt une autre importance.

Deuxièmement, la difficulté rencontrée par Balzer *et al.* (1987) pour prouver l'équivalence des deux formulations n'est pas, en soi, une preuve que la conception sémantique ne parvient pas à se débarrasser des questions de langage, mais doit être interprétée comme le signe soit d'une difficulté formelle importante mais non insurmontable (autrement dit, du fait que la question de leur équivalence n'est pas triviale), soit du fait que ces différences ne sont pas *que* des différences de formulation, et que les structures décrites par les différentes formulations de la mécanique ne sont pas strictement équivalentes. Je voudrais à présent montrer que les seuls outils formels ne suffisent pas à trancher entre ces deux options, et qu'affirmer que les structures décrites par les différentes formulations de la mécanique sont ou non

¹²⁸Moulines est d'ailleurs explicite sur ce point, et admet qu'« il y a dans l'identité essentielle de toute théorie une composante irréductiblement pragmatico-historique, non formalisable » (Moulines, 2006, p. 141). De ce point de vue, ces auteurs n'entrent donc pas dans le champ de ma critique, puisque c'est ce que je vise à démontrer. Je suggère simplement que l'axiomatisation qu'ils proposent de la mécanique, et leur échec – assumé – à rendre compte de son identité sont exemplaires des limites que je cherche à tracer.

isomorphes dépend de facteurs dont l'axiomatisation ne rend pas compte.

On a vu au chapitre 2 (section 4.3) que les formulations newtonienne, lagrangienne et hamiltonienne de la mécanique reposent sur une description de la dynamique des systèmes physiques dans des systèmes de coordonnées différents. En conséquence, la structure géométrique de l'espace dans lequel l'état des systèmes est représenté n'est pas la même selon la formulation adoptée : la formulation hamiltonienne assigne à cet espace une structure symplectique, la formulation lagrangienne une structure métrique, et la formulation newtonienne une structure vectorielle. Ces différences n'impliquent et ne peuvent impliquer de différences empiriques : les contraintes dynamiques imposées par les équations dans ces différentes structures reviennent à décrire la même classe de trajectoires possibles. Cependant, la structure des espaces géométriques sous-jacents aux différentes formulations impose des contraintes différentes sur les mouvements géométriquement possibles. À ce point, si l'on veut trancher, au moyen d'une reconstruction formelle, la question de l'identité de ces formulations, cette dernière doit nous fournir le moyen d'établir un critère de distinction entre les structures qui font effectivement partie du contenu de la théorie, et ce qui n'est qu'un effet de la formulation choisie pour les décrire. Ne peut-on pas envisager de distinguer, dans ce cas précis, entre la structure de l'espace géométrique dans lequel les équations de la mécanique décrivent la trajectoire des systèmes et les modèles destinés à avoir une contrepartie empirique ? Cela reviendrait à dire que les formulations de la mécanique ont le même contenu, quand bien même ce dernier n'est pas décrit au même niveau de généralité. Autrement dit, cela reviendrait à dire que la structure géométrique de l'espace dans lequel elles décrivent la dynamique des systèmes ne fait pas partie de leur contenu.

On le peut. Mais les seuls outils formels de l'axiomatisation ne suffisent pas à soutenir une telle affirmation. Elle dépend, en effet, de ce que l'on entend par « structure », et, corrélativement, de ce que l'on considère être l'objet de la mécanique. Aucune de ces deux choses ne va de soi. Pour le comprendre, revenons un instant à la manière dont Suppes (1968) présente les avantages de l'axiomatisation des théories. Axiomatiser une théorie, selon lui, permet d'en isoler les hypothèses fondamentales ; cela offre les moyens, le cas échéant, de montrer l'équivalence entre formulations différentes. Ainsi, par exemple, la théorie de la relativité restreinte, quelle que soit la manière dont on préfère la formuler, repose sur l'hypothèse selon laquelle « tous les observateurs au repos dans des référentiels inertiels obtiennent les mêmes mesures des distances relativistes le long des trajectoires inertielles quand leurs instruments de mesure sont calibrés de façon identique. » (Suppes, 1968, p. 657)

Selon cette conception, l'axiomatisation des formulations newtonienne, lagrangienne et hamiltonienne devrait permettre d'isoler leurs hypothèses fondamentales

de manière à montrer qu'elles sont, ou ne sont pas, les mêmes. Or, *selon ce qu'on considère être l'objet de la mécanique classique*, on jugera que la structure géométrique assignée par les différentes formulations à l'espace des états fait partie ou non de ses hypothèses fondamentales.

Le principe fondamental de la dynamique, en tant qu'il décrit l'évolution dynamique des systèmes dans un espace vectoriel, affirme, en un sens, que la structure géométrique du monde est vectorielle. Comme je l'ai suggéré au chapitre 2, page 78, reprenant l'argument de Jill North (2009b), une certaine version du réalisme structural pourrait conduire à considérer que la formulation newtonienne assigne *au monde* une structure géométrique différente de celle que les formulations lagrangienne et hamiltonienne, respectivement, lui assignent. En outre, comme je l'ai également suggéré, si l'on conçoit la mécanique non seulement comme la science du mouvement dans l'espace et le temps, mais aussi comme la théorie de l'espace et du temps en tant qu'ils contraignent les mouvements, alors, certainement, les formulations newtonienne, lagrangienne et hamiltonienne ne reposent pas sur les mêmes hypothèses fondamentales. *Si l'on adopte une telle conception*, on affirmera que l'axiomatisation permet de mettre en évidence que certaines sous-structures des différentes formulations sont identiques, mais que leur « superstructure » théorique diffère.

En revanche, *si l'on adopte une conception différente de la mécanique*, et que l'on n'attend d'elle qu'une description de l'évolution dynamique des corps dans l'espace et le temps, on trouvera dans l'axiomatisation des trois formulations le moyen d'en montrer l'équivalence. On pourra alors affirmer que l'axiomatisation permet de mettre en évidence que ces différences ne sont *que* des différences de formulation et n'affectent pas le contenu de la mécanique : selon la manière dont elle est formulée, elle semble dire des choses différentes (puisque la mécanique newtonienne semble dire que l'espace a une structure vectorielle quand la mécanique hamiltonienne semble dire que l'espace a une structure symplectique), mais la reconstruction formelle met en évidence que, en fait, elle dit bien la même chose, quelle que soit sa formulation.

Il ressort de cette analyse la chose suivante : la reconstruction formelle, à elle seule, *ne suffit pas* à prouver que les différentes formulations de la mécanique ne sont que des différences de formulation, ou, au contraire, qu'elles impliquent des différences structurelles. Elle *peut servir à montrer* l'une ou l'autre chose, selon ce que l'on considère être l'objet de la mécanique. Ce dernier élément détermine partiellement la manière même dont on va construire l'axiomatisation de la théorie. C'est ici que ma remarque à propos du choix, par Balzer *et al.* (1987), d'axiomatiser la mécanique newtonienne d'une part et la mécanique lagrangienne de l'autre, prend son sens : c'est parce qu'ils ont une certaine conception de l'objet de la mécanique et de ce qu'elle est censée décrire et expliquer qu'ils affirment, d'emblée, que les mé-

caniques newtonienne et lagrangienne sont deux théories différentes. Cette option – qui n’est, en elle-même, soutenue par aucune axiomatisation préalable – détermine la manière dont ils reconstruisent les deux théories. Certes, le langage dans lequel on choisit d’axiomatiser une théorie est d’importance secondaire, puisque son contenu consiste dans les structures que l’on décrit au moyen de ce langage. Cependant, choisir d’axiomatiser la mécanique *newtonienne*, c’est choisir un langage mathématique qui décrit ces structures à un niveau de généralité différent de celui de la mécanique hamiltonienne. C’est choisir des variables primitives et des lois fondamentales qui incitent à voir dans cette différence de généralité une authentique différence de contenu. Non seulement l’axiomatisation, en tant que telle, ne permet pas de trancher entre différence de formulation et différence de contenu – différence *théorique* –, mais elle est elle-même partiellement déterminée par les engagements épistémiques et les choix des agents qui la construisent et en tirent des conclusions. Autrement dit, pour reprendre une notion introduite dans la première partie de ce travail, elle est elle-même partiellement déterminée par la *version* que les agents ont de la mécanique et de ses formulations.

La conception sémantique, on l’a vu, a pris acte du fait qu’il convenait de se donner tous les outils mathématiques nécessaires à la reconstruction du contenu des théories et que la formalisation stricte envisagée par les empiristes logiques est irréalisable et vaine. Cependant, franchir ce pas – indispensable – interdit d’affirmer que la reconstruction formelle permet de dissocier le contenu d’une théorie de sa formulation. La distinction entre différence de formulation et différence de contenu, dont la reconstruction formelle doit rendre compte, n’est pas du ressort de cette reconstruction elle-même. Selon le *niveau de généralité* auquel on choisit d’axiomatiser la mécanique – selon que l’on commence par axiomatiser séparément les trois formulations, ou qu’au contraire on part d’une axiomatisation de la formulation hamiltonienne pour montrer qu’elle décrit les mêmes structures, à un niveau plus général, que les formulations newtonienne et lagrangienne – et selon la manière dont on choisit d’interpréter ce que l’axiomatisation met en évidence, on jugera ou non que les hypothèses fondamentales de ces formulations sont les mêmes. Autrement dit, le résultat d’une reconstruction formelle (dès lors qu’il ne s’agit pas d’une formalisation stricte) peut être lui-même considéré comme une nouvelle formulation.¹²⁹

Mon propos ne revient pas à nier que l’axiomatisation sémantique des théories

¹²⁹ Comme je l’ai suggéré précédemment (note 126), c’est sans doute parce que les axiomatisations de la conception sémantique peuvent être considérées comme des reformulations des théories que, selon moi, les défenseurs de cette conception confondent parfois la formalisation stricte de la conception syntaxique avec une formulation, ce qui explique les critiques infondées que j’ai signalées ci-dessus (voir section 3.2).

soit un outil puissant d'exploration de leur contenu. Elle permet sans doute de clarifier les relations entre les structures décrites par les différentes formulations de la mécanique. En revanche, ce qui permet de distinguer entre une authentique différence de structure (et donc de contenu) et une simple différence de formulation, inessentielle au contenu de la théorie, n'est pas la méthode axiomatique elle-même. Cette distinction dépend de la manière dont les agents utilisent et comprennent cette théorie. Suppes, tout comme van Fraassen, reconnaît l'importance de certains éléments pragmatiques de l'activité théorique (en particulier pour l'application empirique des théories) ; je suggère pour ma part qu'ils interviennent toujours, *in fine*, quand on cherche à spécifier le contenu d'une théorie par la méthode axiomatique. Il y a donc toujours, dans l'exploration même du contenu des théories, l'intervention d'un aspect essentiellement pragmatique, au sens où il est lié au contexte d'utilisation de la théorie, dont les seuls outils formels ne suffisent pas à rendre compte.

Notons, pour finir, que cela permet d'éclairer rétrospectivement l'usage équivoque que Nagel fait de la notion de « modèle », et sa position ambiguë sur la manière dont il convient de traiter les différentes formulations de la mécanique. Les modèles d'une théorie, on l'a vu, sont portés par ses formulations particulières et les analogies qu'elles suggèrent, qui donnent de la chair à la structure syntaxique. Ils semblent en ceci fort différents des modèles de la conception sémantique, qui sont les structures décrites par les différentes formulations. Cependant, avec la remise en cause que je viens d'opérer d'une distinction nette entre formulation et reconstruction formelle, on peut comprendre que, selon la manière dont on utilise et dont on comprend les différentes formulations de la mécanique, on peut les considérer comme les formulations d'une même théorie ou bien comme des théories différentes, portées par des modèles différents.

Si mon analyse est correcte, on en arrive donc à la conclusion suivante : *selon ce que l'on considère être l'objet de la mécanique et selon l'attitude épistémique que l'on adopte à son égard, la différence entre les formulations de la mécanique peut être traitée comme une pure différence de formulation ou comme une authentique différence de structure. Autrement dit, la frontière entre ce qui relève de la formulation et ce que la reconstruction formelle doit isoler est déterminée par des facteurs dont les outils formels ne rendent pas compte. En conséquence, ils ne suffisent pas, à eux seuls, à la tracer.*

Conclusion

Les approches formelles, qui cherchent à reconstruire le contenu des théories en traitant leur formulation comme inessentielle et, le cas échéant, à dégager l'équiva-

lence structurelle derrière des différences de formulation, se heurtent au problème suivant : il n'est pas possible d'identifier ce contenu sans prendre en compte, dans la reconstruction formelle elle-même, l'utilisation que les agents font des théories. C'est le problème qui a conduit Nagel à intégrer dans les composantes essentielles des théories la notion de « modèle », mettant ainsi en évidence le fait qu'une théorie est toujours déjà interprétée. Ce faisant, il a montré l'importance de certains aspects de la formulation des théories, auxquels je m'intéresse dans le reste de ce travail. La conception sémantique, prenant acte du fait qu'une reconstruction purement syntaxique est impossible, a tenté à son tour d'évacuer les aspects inessentiels des formulations des théories en cherchant dans la notion mathématique de modèle un outil permettant de rendre compte de l'éventuelle équivalence structurelle de différentes formulations. Pourtant, les outils formels à eux seuls ne suffisent pas à prouver cette équivalence. La manière dont les agents utilisent et comprennent les théories – sous telle ou telle formulation particulière – intervient toujours dans l'exploration de leur structure et dans l'affirmation d'une équivalence structurelle entre deux formulations. Il ne s'agit donc pas de dire que l'approche formelle de la conception sémantique est stérile – elle n'a plus besoin d'avocats pour montrer ses succès – mais qu'elle ne peut proclamer, comme elle le fait parfois, s'être débarrassée de certains aspects, jugés inessentiels, liés à la formulation des théories.

Dans la première partie de ce travail, consacrée à la mécanique classique, j'ai montré que la notion de théorie, traditionnellement conçue comme un ensemble d'hypothèses jointes à la totalité de leurs conséquences déductives, ne permet pas de rendre compte dans le détail de l'activité théorique des agents. Dans la mesure où les conceptions syntaxique et sémantique, qui se sont efforcées de donner une caractérisation formelle des théories, ne prétendent pas rendre compte de cette activité, elles ne sont pas mises en défaut par ce constat. En ce sens, mon travail dans son ensemble, qui vise à étudier la manière dont les agents comprennent et manipulent les représentations qu'ils utilisent, ne constitue pas une critique interne des approches formelles ; il propose simplement de rendre compte de ce qu'elles situent explicitement hors du champ de leur analyse. J'espère cependant, dans ce chapitre, avoir montré que ces aspects liés à l'utilisation des théories marquent les limites internes du projet de reconstruction formelle : ce sont eux qui, en dernière analyse, déterminent la limite entre ce qui relève du contenu d'une théorie et ce qui relève de sa formulation.

Dans ce qui suit, je ne poursuis pas l'analyse de l'équivalence et des différences *structurelles* entre formulations de la mécanique. Prenant acte du caractère insuffisant de la notion de théorie pour rendre compte de l'activité théorique, je renonce à en chercher une définition et je propose d'autres unités et outils d'analyse de l'interaction cognitive des agents avec les représentations qu'ils utilisent.

Appendice : La version finale de la conception syntaxique selon Frederick Suppe (1977a, pp. 50-52)

[Les théories scientifiques] ont une formulation canonique qui satisfait les conditions suivantes :

1. Il existe un langage logique de premier ordre L dans les termes duquel la théorie est formulée, et un calcul formel K défini dans les termes de L .
2. Les constantes primitives non logiques ou descriptives (i.e., les « termes ») de L sont séparées en deux classes disjointes :
 - V_O qui contient uniquement les termes d'observation.
 - V_T qui contient les termes non observationnels, ou termes théoriques. V_O doit contenir au moins une constante d'individu.
3. Le langage L est divisé dans les sous-langages suivants, et le calcul formel K est divisé dans les deux sous-calculs formels suivants.
 - (a) Le *langage d'observation*, L_O , est un sous-langage de L qui ne contient ni quantificateurs ni modalités, et qui contient les termes de V_O mais aucun terme de V_T . Le calcul formel K_O qui lui est associé est la restriction de K sur L_O , et doit être tel que tout terme qui n'appartient pas à V_O (i.e., les termes non primitifs) dans L_O est explicitement défini dans K_O ; en outre, K_O doit admettre au moins un modèle fini.
 - (b) Le *langage d'observation logiquement étendu*, L'_O , ne contient pas de termes de V_T et peut être considéré comme étant formé à partir de L_O en y ajoutant les quantificateurs, les modalités, etc., de L . Le calcul formel K'_O qui lui est associé est la restriction de K sur L'_O .
 - (c) Le *langage théorique*, L_T , est le sous-langage de L qui ne contient pas de termes de V_O ; le calcul formel qui lui est associé, K_T , est la restriction de K sur L_T .

L'ensemble de ces sous-langages n'épuisent pas le contenu de L , car L contient aussi des *énoncés mixtes* – ceux dans lesquels apparaissent au moins un terme de V_T et un terme de V_O . De plus, on suppose que chacun des sous-langages ci-dessus a son propre stock de prédicats et / ou de variables fonctionnelles, et que L_O et L'_O ont le même stock, qui est distinct de celui de L_T .

4. L_O et le calcul formel qui lui est associé reçoit une *interprétation sémantique* qui satisfait les conditions suivantes :
 - (a) Le domaine d'interprétation consiste en des événements, des choses ou des choses-instants concrets et observables ; les relations et les propriétés de l'interprétation doivent être directement observables.
 - (b) Toutes les valeurs de chaque variable dans L_O doivent être désignées par une expression de L_O .

Par conséquent, toute interprétation ainsi définie de L_O et K_O , quand elle est augmentée par les règles de vérité appropriées additionnelles, va devenir une interprétation de L'_O et K'_O . On peut considérer les interprétations de L_O et K_O comme des *interprétations sémantiques partielles* de L et K , et on exige que L et K ne reçoivent pas d'autre interprétation sémantique observationnelle que celle fournie par de telles interprétations sémantiques partielles.

5. Une *interprétation partielle* des termes théoriques et des énoncés de L qui les contiennent est fournie par les deux sortes de postulats suivants : les *postulats théoriques* T (i.e., les axiomes de la théorie) dans lesquels seuls des termes de V_T apparaissent, et les *règles de correspondance* ou postulats C , qui sont des énoncés mixtes. Les règles de correspondance C doivent satisfaire les conditions suivantes :
 - (a) L'ensemble des règles C doit être fini.
 - (b) L'ensemble des règles C doit être logiquement compatible avec T .
 - (c) C ne contient pas de terme extra-logique qui n'appartienne ni à V_O ni à V_T .
 - (d) Chaque règle de C doit contenir au moins un terme de V_O et un terme de V_T de manière essentielle et non triviale [*nonvacuously*].

Appelons T la conjonction des postulats théoriques et C la conjonction des règles de correspondance. Alors, la théorie scientifique basée sur L , T , et C consiste dans la conjonction de T et de C et est désignée par TC .

Chapitre 6

Paradigmes, modèles et gabarits : unités d'analyse globales et locales



Camille Flammarion, *Universum*, gravure sur bois pour *L'Astronomie Populaire*, Paris, 1888
(Couleurs : Hugo Heikenwaelder, Vienne, 1998).

L'analyse de l'exemple de la mécanique classique, dans la première partie de ce travail, m'a permis de montrer, entre autres, deux choses dont ce chapitre et le suivant vont chercher à rendre compte. Premièrement, j'ai mis en évidence l'importance de la *forme* des représentations utilisées par les agents : équivalentes *en principe*, les différentes équations de la mécanique classique ne représentent pas les phénomènes de la même manière, et ne permettent pas les mêmes inférences. Deuxièmement, en adoptant le point de vue des agents et en étudiant la manière dont ils comprennent les représentations qu'ils utilisent, j'ai montré que la distinction entre une différence formelle et une différence conceptuelle, et donc entre ce qui constitue le contenu d'une théorie et ce qui lui est inessentiel, n'est pas tranchée ; elle dépend parfois de ce que j'ai appelé la « version » particulière de chaque individu (ou groupe d'individus). Cette version se caractérise, entre autres, par la manière dont il ordonne les différentes représentations (principes, théorèmes, lois, etc.) et dont il les relie aux phénomènes empiriques ; ces deux choses (ordre des principes et lien au monde empirique) déterminent partiellement la signification de ces représentations.

En conséquence, la notion de théorie semble à la fois trop large, trop étroite, et trop abstraite pour rendre compte de l'activité théorique. *Trop large*, parce que les agents, de l'étudiant à l'expert, ont rarement affaire à la théorie entière, mais plutôt à des « morceaux » qu'ils utilisent pour représenter les phénomènes. Qu'il s'agisse de décrire le comportement d'un système physique ou d'explorer les conséquences d'un principe, on raisonne au moyen de représentations particulières.

Trop large, la notion de théorie est aussi, en un autre sens, *trop étroite*. En effet, l'examen de différentes versions de la mécanique (chapitre 3) a montré que la manière dont chacun comprend et ordonne les différents principes et concepts de la mécanique dépend partiellement de facteurs externes à la mécanique. D'une part, elle dépend de facteurs psychologiques (préférences individuelles pour un type de représentation, habitudes de raisonnement, compétences et talents). D'autre part, elle dépend de facteurs contextuels, et tout particulièrement du contexte scientifique, c'est-à-dire de l'ensemble de la connaissance théorique disponible dans la communauté à laquelle appartient l'agent : à partir du moment où l'on décompose la notion de théorie et où l'on prête attention à la forme des représentations et à la manière dont les agents les ordonnent et les utilisent, la frontière entre différentes « théories » se brouille ; en effet, plusieurs représentations, comme par exemple les équations hamiltoniennes, sont communes à différentes « théories ». La manière dont un agent utilise les représentations issues d'« une certaine théorie » est donc partiellement déterminée par l'utilisation qu'il fait de représentations issues de ce que l'on considère généralement comme « une autre théorie ». Une « théorie » est donc rarement abordée comme un tout *isolé* des autres « théories ». C'est en ce sens que l'on peut dire que cette notion

est tout à la fois trop large et trop étroite.

Enfin, la notion de théorie est *trop abstraite*. Non seulement les agents ont affaire à différentes représentations diversement liées entre elles et non à un tout isolé, mais encore la signification de ces représentations, leur dimension *conceptuelle*, dépend du fait qu'elles sont *comprises* par des agents. Comme le disent Griesemer et Wimsatt (1989), « les théories exigent des théoriciens ».¹ Il ne s'agit donc pas d'étudier la relation entre un ensemble de principes abstraits et les phénomènes qu'ils sont censés représenter, mais l'*interaction cognitive des agents avec ces représentations*. Étudier cette interaction cognitive implique de prêter attention à la fois à la forme particulière des représentations et à la manière dont les agents les relient au monde, et entre elles.

Dans ce chapitre, je souhaite revenir sur certaines analyses de l'activité théorique qui ont été proposées par des critiques des positivistes logiques et, plus récemment, des critiques de la conception sémantique des théories (et des approches formelles en général). Chacune de ces analyses consiste, entre autres, à définir de nouvelles notions destinées à remplacer – ou, selon les cas, à amender ou compléter – celle de théorie. Ces notions – *paradigmes* (Kuhn, 1962/1970), *modèles* (Giere, 1988; Cartwright, 1983, 1999; Morgan et Morrison, 1999; Godfrey-Smith, 2006; Frigg, 2006)², *gabarits* (Humphreys, 2004) – visent à isoler dans la pratique scientifique des éléments dont la description et l'analyse doivent permettre de rendre compte de la connaissance produite par l'entreprise scientifique. C'est ce que j'appelle, après Paul Humphreys (2004, pp. 57-60), des « *unités d'analyse* » de l'activité théorique.

L'introduction de ces nouvelles unités d'analyse – et l'abandon, l'amendement ou l'enrichissement de celle de théorie – sont la plupart du temps motivés par un constat au moins partiellement similaire à celui que je viens d'énoncer concernant les inconvénients et les limitations de la notion de théorie. Je me propose donc, dans ce qui suit, d'examiner certaines de ces unités d'analyse et de les évaluer à la lumière des exigences qui se dégagent, en négatif, des inconvénients de la notion de théorie que j'ai pointés : ces unités d'analyse doivent permettre de rendre compte à la fois du fait que les agents utilisent des représentations sous une forme particulière, des facteurs

¹ « [...] de nombreux philosophes des sciences parlent des théories comme d'entités abstraites, mais les théories exigent des théoriciens, et les entités abstraites, des incarnations. » [*theories require theorizers, and abstract entities, entifiers*] (Griesemer et Wimsatt, 1989, p. 87)

² Je ne mentionne ici que (quelques) analyses récentes. L'importance de la notion de modèle avait déjà été soulignée dans les années 1960, comme on l'a vu au chapitre 5, par Nagel (1961), ainsi que par Hesse (1966) et Black (1962). Elle est aussi, on l'a vu, au cœur de la conception sémantique des théories ; cependant, les analyses auxquelles je m'intéresse ici soulignent l'importance des modèles d'un point de vue pratique et cognitif, et sont le plus souvent des critiques de la conception sémantique – à l'exception toutefois de celle de Giere (1988), qui tente de montrer la compatibilité de la conception sémantique avec une approche cognitive.

psychologiques et contextuels qui influent sur la manière dont ils les utilisent, et du fait que leur signification dépend partiellement de la manière dont ils les utilisent et, de ce fait, les comprennent.

Dans un premier temps (section 1), je présenterai les avantages et les limites d'une approche privilégiant des unités d'analyse *globales*, c'est-à-dire non seulement plus larges que celle de théorie, mais aussi plus hétérogènes ; la notion de paradigme (Kuhn, 1962/1970), directement orientée contre celle de théorie telle que les empiristes logiques la conçoivent, en est l'exemple le plus fameux. Au moyen de cette notion, Kuhn cherche à rendre compte à la fois de l'importance du contexte scientifique – au sens large où il inclut les pratiques expérimentales, les institutions scientifiques, ainsi que les croyances métaphysiques d'une époque –, mais aussi de la dimension pratique et cognitive de l'activité scientifique ; on verra en effet qu'il prête une attention particulière à l'apprentissage scientifique et aux pratiques concrètes par lesquelles les scientifiques donnent une signification aux représentations qu'ils utilisent. Je montrerai cependant que la notion de paradigme souffre d'une double limitation. D'une part, son caractère hétérogène et très englobant la rendent bien souvent inopérante. D'autre part, la thèse kuhnienne de l'incommensurabilité des paradigmes conduit à identifier ces derniers à des cadres linguistiques et à admettre une conception logico-empirique de leur contenu ; ils ne se distinguent alors plus très nettement des théories.

Prenant acte de la nécessité de recentrer l'analyse sur des unités plus *locales*, j'examinerai rapidement (section 2) la notion de modèle. S'intéresser aux modèles semble être un bon moyen de tirer les fruits de l'analyse kuhnienne tout en se débarrassant des aspects globalistes et linguistiques de son approche. Je montrerai cependant que les analyses philosophiques contemporaines de cette notion restent souvent tributaires d'une conception trop abstraite de la représentation scientifique, et n'accordent pas assez d'importance à la forme particulière de ces représentations pour les raisonnements des agents qui les utilisent.

Pour finir (section 3), je reviendrai sur une des composantes de la notion de paradigme, que Kuhn (1969, 1970a,b) appelle « cas exemplaires », et qui permet de soulever plusieurs questions importantes sur la manière dont, en pratique, les agents relient les représentations particulières qu'ils utilisent aux phénomènes. Je procéderai à un examen critique de ces questions à la lumière des réflexions de Paul Humphreys (2004) sur ce qu'il appelle les « gabarits » ; au moyen de cette notion, Humphreys souligne l'importance de la forme particulière – ce qu'il appelle la « syntaxe » – des « morceaux » de théorie que l'on utilise pour prédire et expliquer les phénomènes. Cela me permettra de préciser la relation entre les deux dimensions – « computationnelle » et « représentationnelle » – de la compréhension que j'ai mentionnées précédemment (chapitre 4), et me conduira à conclure que, dès lors que l'on prête attention à la forme

concrète des représentations utilisées dans la pratique scientifique, il n'y a plus aucune raison de se concentrer uniquement sur les représentations linguistiques (comme les équations), et qu'il convient de prêter attention à *tous* les types de représentations.

1 Motivations et limites d'une approche globaliste : la notion de paradigme (Kuhn, 1962/1970)

Un des inconvénients de la notion de théorie pointés dans l'introduction de ce chapitre est ce que j'ai appelé son caractère trop *étroit* : en prêtant une attention exclusive à ce que l'on appelle couramment des « théories » – la mécanique classique, la théorie de l'évolution, les théories de la relativité restreinte et générale, etc. – et en réduisant leur signification, comme le font les positivistes logiques, à leur contenu empirique, on néglige l'importance du contexte dans lequel ces théories sont formulées et utilisées. Pour prendre un exemple simple, ignorer le contexte scientifique et linguistique dans lequel une théorie est comprise par les agents revient à négliger l'importance du fait que les termes employés dans la formulation de ses hypothèses (comme celui de force en mécanique) ont bien souvent un sens préalable, soit qu'ils appartiennent au langage ordinaire, soit qu'ils apparaissent dans d'autres théories.

Afin de rendre compte de l'importance des déterminants contextuels de l'activité scientifique, plusieurs historiens et philosophes des sciences ont proposé, autour des années 1960, de s'intéresser à des unités plus larges, et souvent beaucoup plus hétérogènes, que celles de théorie.³ Celles-ci reçoivent des noms et des définitions différentes selon les auteurs.⁴ Dans ce qui suit, je m'intéresse exclusivement à la plus

³Oltre Thomas Kuhn, les plus célèbres sont Paul Feyerabend (1962, 1965b), Norwood Russell Hanson (1958) et Stephen Toulmin (1959, 1961). Dans les années 1970, les travaux de Imre Lakatos (1970) et de Larry Laudan (1977) se situent également dans cette perspective. Ce renouveau en philosophie des sciences a été baptisé « révolte contre le positivisme » (Shapere, 1966, p. 41), en raison du fait qu'il repose sur un rejet en bloc de la plupart des présupposés sur lesquels est fondée l'entreprise formaliste des positivistes logiques. Ces analyses ne sont cependant pas exclusivement dirigées contre l'empirisme logique, que leurs défenseurs connaissent parfois assez mal. Elles s'opposent, plus généralement, à une image des sciences positiviste au sens large – dont les empiristes logiques offrent une version formelle et hautement sophistiquée – selon laquelle l'histoire des sciences est celle d'un progrès cumulatif dont le moteur principal est l'expérience. Voir (Shapere, 1966; Hacking, 1981; Laudan *et al.*, 1986) pour une analyse des présupposés communs et des différences entre ces critiques du positivisme.

⁴Paul Feyerabend (1962, 1965b) les appelle simplement « théories », mais donne à ce terme une signification beaucoup plus englobante que celle de la notion classique de théorie. Stephen Toulmin parle d'« idéaux d'ordre naturel » et les définit comme des « standards de rationalité et d'intelligibilité », au travers desquels « nous voyons le monde [...] au point que nous oublions de quoi il aurait l'air sans eux » (Toulmin, 1961, p. 101), et qui « donnent une signification aux [faits]

célèbre d'entre elles, celle de *paradigme*, que Thomas Kuhn (1962/1970) introduit dans *La Structure des révolutions scientifiques* et sur laquelle il reviendra tout au long de sa vie, en réponse aux critiques qu'elle n'a cessé de susciter.

La notion de paradigme étant hautement polysémique⁵, il est difficile d'en donner une définition simple qui ne soit pas du même coup très partielle. Disons pour commencer qu'un paradigme est un ensemble de présupposés souvent implicites de différents ordres (scientifiques, méthodologiques, métaphysiques, etc.), partagés par une communauté scientifique. Selon Kuhn, l'entreprise scientifique repose sur l'adoption préalable de cet ensemble de présupposés, qui détermine les questions susceptibles de recevoir une explication scientifique, les méthodes qu'il convient de mettre en ordre pour y répondre, et les réponses dont on considérera qu'elles sont effectivement explicatives.

Cela ne revient pas seulement à affirmer que l'entreprise scientifique, comme toute entreprise humaine, est déterminée contextuellement, mais encore qu'il existe un cadre conceptuel *plus large* que les théories scientifiques définies comme des ensembles d'hypothèses confrontées à l'expérience, *cadre qui détermine au moins partiellement le contenu des produits de l'activité scientifique*.⁶ Contrairement aux théories des empiristes logiques, ce cadre conceptuel, souvent décrit comme une « vision du monde »⁷, n'est ni confirmé ni infirmé par l'expérience ; il est ce qui rend la possible, allant jusqu'à déterminer, selon certains propos de Kuhn inspirés de Hanson (1958), l'observation elle-même.⁸

En conséquence, selon Kuhn, le contenu des produits de l'activité scientifique n'est pas réductible à un contenu empirique indépendant du cadre dans lequel il est appréhendé et de la forme sous laquelle il est présenté ; corrélativement, comprendre

et déterminent même ce qui compte comme des « faits » pour nous » (p. 95) ; parmi ces faits, ils marquent ceux qui, par contraste avec la « course naturelle des événements » (p. 79), appellent une explication scientifique. Imre Lakatos (1970) parlera pour sa part de « programmes de recherches » et Larry Laudan (1977) de « traditions de recherche ».

⁵Margaret Masterman (1970) dresse ainsi la liste des 22 sens différents sous lesquels la notion apparaît dans *La Structure des révolutions scientifiques*.

⁶Contrairement aux catégories kantienne, ce cadre conceptuel *a priori* est éminemment historique et contextuel. L'approche kuhnienne, ainsi que celle des autres acteurs de la « révolte contre le positivisme » (voir note 3), est souvent dite « historiciste », en raison du fait qu'elle affirme que la science est un phénomène *intrinsèquement* historique, et qu'une analyse des produits de l'activité scientifique qui ne rend pas compte de cette historicité manque une part importante de leur contenu.

⁷On appelle parfois les travaux de la tradition critique de l'empirisme logique dans laquelle Kuhn s'insère des « analyses de la *Weltanschauung* » (Suppe, 1977a, pp. 125-221).

⁸C'est ce que signifie l'affirmation selon laquelle l'observation est « chargée de théorie » [*theory-laden*]. Ici, la « théorie » qui imprègne l'observation n'est pas le système déductif des empiristes logiques, mais bien l'ensemble des présupposés, explicites ou implicites, d'une communauté.

une « théorie » ne revient pas à identifier la signification empirique d'un ensemble d'énoncés, mais à maîtriser un cadre conceptuel, dont l'acquisition ne s'apparente pas à l'apprentissage de « principes » et de « lois » mais plutôt à celui d'une langue naturelle. C'est de ce cadre conceptuel que l'expérience et les théories qui rendent compte de cette expérience tirent leur signification et leur intelligibilité.

Deux présupposés majeurs de l'œuvre de Kuhn vont m'intéresser dans ce qui suit. Le premier est l'idée selon laquelle une analyse de la connaissance scientifique doit s'intéresser en priorité à deux aspects de l'activité scientifique : d'une part, au processus d'acquisition de cette connaissance, c'est-à-dire à l'*apprentissage scientifique*, dont Kuhn considère que l'analyse fournit la clé de celle de la connaissance scientifique en général ; d'autre part, aux *pratiques* concrètes par lesquelles passe cet apprentissage et, plus généralement, par lesquelles on applique et développe la connaissance théorique. On verra, à la section 1.1, que la notion de paradigme, par l'importance qu'elle accorde à l'apprentissage et à la pratique, permet ainsi d'éclairer plusieurs des éléments qui déterminent ce que j'ai appelé, au chapitre 3, les « versions » des agents, c'est-à-dire la manière dont ils comprennent les représentations théoriques auxquelles ils sont confrontés.

Le second présupposé de Kuhn, lié au précédent, est aussi celui qui marque les limites de son approche. L'idée selon laquelle l'entreprise scientifique est rendue possible par l'adoption préalable d'un ensemble de présupposés – par l'existence d'un paradigme – le conduit, comme on le verra, à tracer une frontière imperméable entre la pratique de ce qu'il appelle la *science normale* et les changements qu'il décrit comme *révolutionnaires*. Cette distinction radicale entre deux types de changements scientifiques, et la volonté de rendre compte des révolutions scientifiques comme des changements de paradigmes, le conduisent, comme on le verra à la section 1.2, à figer le paradigme en une unité monolithique uniquement définie en termes linguistiques, gommant finalement la différence entre paradigmes et théories.

1.1 Le paradigme comme ciment de la communauté scientifique

L'apprentissage scientifique, au cours duquel un étudiant devient progressivement membre d'une communauté scientifique et praticien d'une discipline, consiste essentiellement, selon Kuhn, en l'acquisition d'un ensemble de méthodes et de pratiques, et, du même coup, en l'adoption progressive des présupposés partagés par les membres de la communauté scientifique en question. C'est cet ensemble hétérogène d'engagements métaphysiques, scientifiques, méthodologiques, et de techniques expérimentales que Kuhn, en 1962, appelle « paradigme ». L'adoption de ces présupposés et la maîtrise de ces techniques sont les déterminants de l'appartenance à une

communauté scientifique et les préalables indispensables à l'exercice de ce que Kuhn appelle la « science normale ».

La science normale⁹ (voir Kuhn, 1962/1970, chap. 3), c'est-à-dire la science institutionnelle, celle qui est enseignée dans les manuels et pratiquée dans les laboratoires de recherche, consiste essentiellement, selon Kuhn, à renforcer et à développer le paradigme qui constitue le ciment de la communauté scientifique ; en conséquence, celui-ci n'est jamais lui-même remis en cause au cours du développement normal d'une science. Bien au contraire, ce « réseau serré d'engagements – conceptuels, théoriques, instrumentaux et méthodologiques » (Kuhn, 1962/1970, p. 42) est la « source des méthodes, des ensembles de problèmes, et des standards de solution acceptés par une communauté scientifique arrivée à maturité à une époque donnée » (Kuhn, 1962/1970, p. 103).

Les paradigmes sont donc des unités bien plus larges que les théories¹⁰ au sens restreint que lui donnent les philosophes des sciences ; ils correspondent cependant, comme le remarque Kuhn, à un sens plus lâche de la notion de théorie, alors proche de celui de l'expression « vision du monde » :

Les scientifiques eux-mêmes diraient qu'ils partagent une théorie ou un ensemble de théories, et j'aimerais que ce terme puisse finalement être réutilisé dans ce sens là. Tel qu'il est aujourd'hui utilisé en philosophie des sciences, cependant, le terme « théorie » désigne une structure dont la nature et la portée sont bien plus limitées que celle dont il s'agit ici. Jusqu'à ce que ce terme soit libéré de ses implications actuelles, on évitera la confusion en en adoptant un autre. (Kuhn, 1970a, p. 182)

Non seulement la notion de paradigme désigne une entité plus large que celle de théorie – puisqu'elle peut inclure plusieurs « théories », ainsi que des croyances non scientifiques – mais elle caractérise en outre une réalité d'une tout autre nature : les paradigmes ne sont pas des corps de connaissances pouvant faire l'objet d'une présentation explicite ; la notion de paradigme désigne plutôt le cadre dans lequel les individus d'une communauté appréhendent un ensemble d'hypothèses théoriques et mettent en œuvre des pratiques expérimentales. En conséquence, la signification des hypothèses théoriques que les membres d'une communauté scientifique énoncent et utilisent pour expliquer certains phénomènes – ce que l'on appelle couramment

⁹Dans cette expression, l'adjectif « normal » est à comprendre principalement au sens de « normatif » : la science normale est celle au sein de laquelle les normes (de rationalité, d'intelligibilité, de méthode, etc.) sont fixées. C'est aussi la science « normée », au sens où elle se définit par le respect de ces normes.

¹⁰Il arrive cependant qu'une théorie, comme par exemple la mécanique classique, incarne à elle seule un paradigme, en devenant un standard d'intelligibilité indiscuté et un modèle pour les autres sciences.

les « théories » – est déterminée par le paradigme dans lequel elles sont énoncées et utilisées.

En substituant la notion de paradigme à celle de théorie, Kuhn affirme donc la chose suivante : le contenu des produits de l'activité scientifique dépend de la manière dont les agents les comprennent. Cette compréhension dépend elle-même de tout un ensemble d'éléments contextuels, qui vont de l'échelle très large de l'époque et de la civilisation auxquelles les agents appartiennent à « l'expérience passée dans d'autres domaines » et au « propre caractère [*makeup*] individuel » de chacun (Kuhn, 1962/1970, pp. 3-4).

Pour étudier ces éléments contextuels, Kuhn choisit l'échelle intermédiaire de la *communauté scientifique*, qu'il définit de façon informelle comme constituée de l'ensemble des « praticiens d'une même spécialité scientifique » liés par un apprentissage scientifique commun, et qui « se considèrent eux-mêmes et sont considérés par les autres comme des hommes poursuivant un ensemble de buts communs, parmi lesquels la formation de leurs successeurs. » (Kuhn, 1969, p. 461)¹¹ Étudier les pratiques qui caractérisent cet apprentissage et la poursuite de ces buts est le moyen, selon Kuhn, de rendre compte de la connaissance scientifique, en tant qu'elle est déterminée par un paradigme.

Il y a ici une difficulté, qu'il convient de souligner dès à présent : comme Kuhn (1970a, 1969) le reconnaît, définir les paradigmes comme ce qui cimente les communautés scientifiques, elles-mêmes étant caractérisées par l'existence d'un paradigme commun, fait courir le risque d'une certaine circularité. Kuhn affirme cependant que la communauté scientifique peut être identifiée au préalable au moyen de critères sociologiques et institutionnels, comme la formation suivie par ses membres, l'existence de revues spécialisées, l'organisation de conférences, la création de chaires universitaires consacrées à une discipline, etc.¹² Cela signifie que ces critères ne constituent pas, en eux-mêmes, le paradigme, mais sont à la fois la cause et le signe de l'existence d'un paradigme commun – défini comme un ensemble d'« engagements ». Autrement dit, le paradigme, selon Kuhn, ne se réduit pas à cet ensemble de réalités sociologiques et institutionnelles : ce terme ne désigne donc pas une unité d'analyse sociologique mais

¹¹ Le texte de Kuhn (1969), intitulé « Second thoughts on paradigms », est celui de son intervention à la conférence sur la structure des théories scientifiques qui s'est tenue dans l'Urbana en 1969. Les actes de cette conférence ont été édités par Frederick Suppe (1974/1977b). La pagination indiquée correspond à cette édition.

¹² Kuhn ne prétend pas lui-même procéder à une caractérisation sociologique des communautés scientifiques, mais affirme que la sociologie des sciences – dans laquelle il ne considère pas que son travail s'insère – devrait pouvoir élaborer les outils nécessaires à une telle caractérisation rigoureuse. Pour toutes ces questions, voir (Kuhn, 1969, pp. 460-462).

plutôt une entité conceptuelle hypothétique censée constituer le socle des pratiques et des croyances des membres de la communauté. Outre que la frontière entre ce qui définit, au niveau sociologique, la communauté, et les engagements que ses membres partagent et qui constituent le paradigme n'est pas très nette, il reste que l'étude des paradigmes passe par celle de réalités sociologiques, pratiques et institutionnelles. Supposer, comme semble le faire Kuhn, que ces différents éléments, fort hétérogènes, sont pour ainsi dire la « partie émergée de l'iceberg » qu'est le paradigme, défini comme une réalité sous-jacente et, en tant que telle, inaccessible à l'analyste, c'est faire un pas de plus. On verra, au terme de cette analyse (section 1.3.1), que c'est cette supposition qui conduit Kuhn à faire du paradigme, conçu de façon globale, une unité monolithique inopérante.

Auparavant, je me propose de montrer que l'examen des différents éléments que Kuhn considère comme les composantes principales du paradigme permet de caractériser certains aspects importants de l'activité théorique. La grande hétérogénéité des composantes du paradigme, si elle en condamne l'unité – dont le maintien conduit aux difficultés que je présenterai à la section 1.2 –, permet en particulier de montrer que plusieurs agents utilisant la même « théorie » peuvent la comprendre différemment – en avoir différentes versions.

1.1.1 Les composantes principales de la « matrice disciplinaire »

En réponse aux nombreuses critiques qui ont pointé la très grande polysémie de la notion de paradigme telle qu'il l'emploie en 1962, Kuhn (1970a, 1969) propose de distinguer entre un sens global et un sens local de cette notion. En son sens global, le paradigme, qu'il rebaptise alors « *matrice disciplinaire* », est défini comme la « constellation entière de croyances, de valeurs, de techniques, etc., partagées par les membres d'une communauté scientifique ». En son sens local, le terme de paradigme désigne une composante particulièrement importante de la matrice disciplinaire, à savoir les « *cas exemplaires* » [*exemplars*], c'est-à-dire les « solutions concrètes de problèmes [...] employées comme modèles ou exemples » dans la formation des étudiants destinés à faire partie d'une communauté scientifique (Kuhn, 1970a, p. 175). En plus des cas exemplaires, Kuhn dégage trois autres composantes de la matrice disciplinaire : les *généralisations symboliques*, les *modèles*, et les *valeurs*. L'analyse de ces différents éléments éclaire, comme on va le voir, la manière dont les agents comprennent les représentations théoriques qu'ils utilisent.

Les généralisations symboliques sont les éléments de la matrice disciplinaire qui semblent se rapprocher le plus des lois scientifiques. Ce sont en effet les « compo-

santes formelles ou facilement formalisables de la matrice disciplinaire » (Kuhn, 1969, p. 463). Elles sont souvent exprimées, surtout en physique, sous la forme d'équations, comme « $f = ma$ » ou « $I = V/R$ » ; elles peuvent aussi être exprimées au moyen d'énoncés du langage naturel, comme par exemple « les éléments se combinent dans des rapports de poids constants » ou « l'action égale la réaction » (Kuhn, 1970a, p. 183).

Ce qui les caractérise en premier lieu, c'est qu'elles sont utilisées « de manière routinière », sans que les scientifiques ressentent « le besoin d'une quelconque justification » (Kuhn, 1969, p. 464). Cependant, si l'utilisation de généralisations symboliques est une condition *sine qua non* de l'appartenance à un groupe scientifique, cette utilisation routinière n'est la marque d'aucun engagement réel de la part de celui qui les utilise car elles n'ont, en elles-mêmes, aucune signification. Elles sont des formes vides, des « esquisses de lois » ou « schémas de lois » (Kuhn, 1970a, p. 188) et ne disent strictement rien sur le monde tant qu'elles ne sont pas appliquées à la résolution de problèmes concrets, sous la forme de cas exemplaires.

Les cas exemplaires [*exemplars*] – qui correspondent au sens local de la notion de paradigme – sont des « solutions concrètes de problèmes que les étudiants rencontrent depuis le début de leur formation scientifique, que ce soit dans les laboratoires, au moment de leurs examens, ou à la fin des chapitres de manuels » (Kuhn, 1970a, p. 187). Typiquement, la résolution du problème du pendule simple, présentée au chapitre 2 (pages 47-48), est un cas exemplaire. C'est par l'application de la généralisation symbolique « $f = ma$ » à la résolution de ce problème que l'étudiant apprend sa signification. C'est par des applications concrètes de ce type qu'elle devient une authentique représentation et que les symboles qui la composent deviennent des concepts. C'est seulement alors qu'elle est *comprise*, et qu'elle peut faire l'objet d'une adhésion de la part de l'agent.

Une des thèses les plus importantes de Kuhn, sur laquelle il insiste dès 1962 et revient avec force dans les années 1970, est que l'application des généralisations symboliques à des problèmes concrets, dont la maîtrise est le déterminant le plus important de l'appartenance à une communauté scientifique, ne consiste pas en l'application de *règles* (voir, par exemple, Kuhn, 1962/1970, p. 42). Si l'on peut toujours, *a posteriori*, énoncer des règles qui permettent de relier les différents symboles de la généralisation symbolique à des situations concrètes, cela ne rend pas compte de la signification qu'elles ont pour les agents. En effet, cette signification, selon Kuhn, est indissociable de la manipulation concrète qu'exige leur application.

En conséquence, la connaissance scientifique n'est pas contenue dans un ensemble de généralisations symboliques jointes à leurs règles d'applications, mais dans les cas

exemplaires, c'est-à-dire les généralisations symboliques *en tant qu'elles sont comprises*. Le cas exemplaire, en effet, ne désigne pas le système concret – ou le système idéalisé (comme par exemple le pendule simple) – auquel est appliquée la généralisation symbolique, mais bien la forme particulière que cette dernière prend quand elle lui est appliquée et acquiert, de ce fait, un contenu. La connaissance scientifique, en ce sens, comporte une grande part de *savoir-faire* – que Kuhn, citant Polanyi (1958), appelle « connaissance tacite » (Kuhn, 1970a, p. 191) – et ne se réduit pas à une connaissance propositionnelle.

Décrire l'application des généralisations symboliques aux situations concrètes comme la mise en œuvre de règles rendrait incompréhensible, selon Kuhn, la capacité des agents qui en maîtrisent l'usage à les écrire sous une forme parfois inédite, adaptée à un nouveau type de problème (Kuhn, 1969, p. 466). Comprendre une généralisation symbolique, c'est-à-dire être en mesure de l'écrire sous la forme adéquate, implique donc, comme je l'ai souligné au chapitre 4 (section 1.2), une double compétence, indissociablement *computationnelle* et *représentationnelle*.¹³ Cela signifie que la capacité à manipuler correctement la généralisation « $f = ma$ » repose sur la capacité à « reconnaître les forces, les masses et les contraintes dans un nouveau problème », ce qui permet alors « d'écrire un formalisme adéquat pour sa solution. » (Kuhn, 1969, p. 477) Cette capacité repose sur celle de « voir différentes situations physiques comme similaires ; ces situations sont, si l'on veut, vues dans une *gestalt newtonienne* ».

Une fois que les étudiants ont acquis la capacité de voir un certain nombre de situations problématiques de cette manière, ils peuvent improviser l'écriture des formes symboliques exigées par d'autres situations qui se présentent. Avant cette acquisition, cependant, la deuxième loi de Newton n'est pour eux rien d'autre qu'une suite de symboles non interprétés. Bien qu'ils la partagent, ils ne savent pas ce qu'elle signifie et elle leur dit par conséquent peu de choses sur la nature. Ce qui leur reste à apprendre n'est pas, cependant, contenu dans des formulations symboliques additionnelles. Au contraire, ils l'acquièrent par le biais d'un processus analogue à celui de l'ostension, la confrontation directe à des séries de situations dont on leur dit que chacune est une situation newtonienne. (Kuhn, 1970b, pp. 169-170)

La connaissance scientifique – la compréhension des hypothèses théoriques – est donc littéralement une pratique, et ne peut pas être analysée par la considération de ces hypothèses théoriques indépendamment de leur utilisation par des agents. Les « lois » scientifiques ne contiennent cette connaissance qu'en tant qu'elles sont utilisées – et

¹³L'analyse de l'articulation de ces deux compétences est l'objet principal de la section 3 du présent chapitre.

comprises – par des agents.

En conséquence, puisque les généralisations symboliques acquièrent leur signification au moyen des situations concrètes auxquelles elles sont appliquées, il est possible que deux agents (ou deux groupes d'agents) utilisent la même généralisation symbolique mais ne la comprennent pas de la même manière, parce qu'ils ne l'appliquent pas aux mêmes cas. Si tous les étudiants en physique apprennent à appliquer « $f = ma$ » à des problèmes basiques comme le mouvement du pendule simple ou la chute libre d'un corps, les cas auxquels ils sont confrontés au fur et à mesure de leur spécialisation diffèrent selon le domaine vers lequel ils s'orientent. De même, bien que « les spécialistes de la physique du solide et de la théorie des champs partagent l'équation de Schrödinger, c'est seulement les applications les plus élémentaires qui sont vraiment communes aux deux groupes » (Kuhn, 1962/1970, p. 187). En conséquence, cette équation n'a *pas exactement* la même signification pour ces deux groupes. Cela ne veut pas dire que sa signification est radicalement différente, mais que, *dans certains cas*, les agents ne l'utilisent pas et, donc, ne la comprennent pas de la même manière.

Ces considérations permettent de dégager des sous-groupes, au sein d'une large communauté, sur la base de la manière dont ils utilisent certaines généralisations symboliques (c'est-à-dire sur la base des cas exemplaires auxquels ils les appliquent).

Sur le chemin de leur spécialisation professionnelle, certains physiciens ne rencontrent que les principes de base de la mécanique quantique. D'autres étudient en détail les applications paradigmatiques de ces principes à la chimie, d'autres encore à la physique des solides, et ainsi de suite. *Ce que signifie la mécanique quantique pour chacun d'eux*¹⁴ dépend de l'enseignement qu'il a suivi, des textes qu'il a lus, et des revues qu'il fréquente. Il s'ensuit que, bien que le changement d'une loi de la mécanique quantique soit révolutionnaire pour tous ces groupes, un changement qui n'a d'incidence que sur l'une ou l'autre application paradigmatique de la mécanique quantique n'est nécessairement révolutionnaire que pour les membres d'une sous-spécialité professionnelle particulière. [...] En bref, bien que la mécanique quantique (ou la dynamique newtonienne, ou la théorie électromagnétique) soit un paradigme pour de nombreux groupes scientifiques, *ce n'est pas le même paradigme pour tous*.¹⁵ En conséquence, il peut déterminer simultanément plusieurs traditions de science normale qui se recoupent sans être coextensives. (Kuhn, 1962/1970, pp. 49-50)

À la limite, on peut dire que chaque individu a son propre paradigme, fruit de son histoire personnelle (de l'enseignement qu'il a suivi, des livres qu'il a lus, etc.). Ce paradigme détermine ce que j'ai appelé, au chapitre 3, sa *version* de la théorie –

¹⁴ Je souligne.

¹⁵ Je souligne.

en l'occurrence, de la mécanique quantique. Sur la base d'un certain nombre de cas exemplaires communs, on peut dégager des groupes d'individus qui entretiennent des versions plus ou moins similaires, à plusieurs niveaux de généralité différents.

Cependant, il est important de noter – et j'y reviendrai plus loin – que cela n'empêche pas les différents individus de s'accorder, dans une plus ou moins grande mesure, sur la signification des généralisations symboliques qu'ils utilisent. Il existe en effet des cas exemplaires communs aux différents sous-groupes d'une communauté « large » : par exemple, tous les utilisateurs de la généralisation symbolique « $f = ma$ » ont appris, au début de leur apprentissage, à l'appliquer au cas du pendule simple. En outre, l'utilisation de cette généralisation symbolique est contrainte par des règles strictes, quand bien même ces règles sous-déterminent la manière dont les agents, confrontés à différentes situations, l'utilisent. Insistons : ces propos de Kuhn n'impliquent pas une conclusion relativiste¹⁶ selon laquelle chacun comprend « ce qu'il veut ». L'activité scientifique est hautement normée, et les différentes applications des généralisations symboliques ne peuvent pas, dans le cadre de la science normale, conduire à des contradictions.

L'analyse kuhnienne permet ainsi de tracer différents types de frontières entre groupes d'individus, et la notion de paradigme semble être d'une très grande plasticité. Les deux autres composantes principales de la matrice disciplinaire (le paradigme au sens global) permettent encore d'affiner l'analyse de certains aspects de l'activité théorique.

Les modèles sont « les engagements de plus haut niveau, quasi-métaphysiques, que l'enquête historique met régulièrement en évidence » (Kuhn, 1962/1970, p. 41). Ce sont les engagements fondamentaux d'un groupe, qui déterminent les méthodes et les standards d'intelligibilité d'une science. Ainsi, « la chaleur est l'énergie cinétique des parties qui constituent les corps » ou « tous les phénomènes perceptibles sont dus à des interactions d'atomes qualitativement neutres dans le vide » (Kuhn, 1970a, p. 184) ne sont pas, selon Kuhn, des lois de la nature ou des généralisations symboliques, mais plutôt des croyances fondamentales qui, sans conduire par elles-mêmes à des prédictions, fournissent le cadre dans lequel l'enquête scientifique peut définir les problèmes à résoudre et la forme des explications à apporter.

Pourquoi Kuhn choisit-il d'appeler ces croyances fondamentales des « modèles », et quel rôle jouent-elles dans la pratique de la science et dans la compréhension des agents ? Un premier élément de réponse nous est fourni par l'affirmation de Kuhn

¹⁶Si Kuhn tient des propos relativistes, c'est bien plutôt, paradoxalement, quand il s'éloigne du détail de la pratique scientifique et accepte plusieurs présupposés des approches formelles. Je montrerai cela à la section 1.2 du présent chapitre.

selon laquelle les modèles « fournissent les analogies et métaphores préférées ou autorisées du groupe. Ce faisant, ils aident à déterminer ce qui sera accepté comme une explication » (Kuhn, 1970a, p. 184).

Prenons, par exemple, l'idée selon laquelle « les molécules d'un gaz se comportent comme de petites boules de billard élastiques en mouvement aléatoire ». L'adoption de ce modèle implique qu'une explication du mouvement des molécules d'un gaz *doit avoir la forme d'une équation de la mécanique classique*. En conséquence, en imposant un type d'explication, le modèle impose le choix d'un certain type de généralisation symbolique. Ce rôle du modèle est à la fois heuristique, puisqu'il aide à trouver la forme des solutions, et normatif. Son adoption signifie qu'une explication qui ne revêtirait pas cette forme ne serait ni vraie ni fausse, mais serait tout simplement non scientifique, voire inintelligible.

Que signifie, à présent, l'affirmation selon laquelle ces modèles sont des « engagements » ? À quel type de croyances les agents qui les acceptent sont-ils engagés ? Sur ce point, Kuhn distingue les modèles qui font l'objet d'un engagement « métaphysique » de ceux qui font seulement l'objet d'un engagement « méthodologique ».

Après 1630, par exemple, et particulièrement après les écrits remarquablement influents de Descartes, la plupart des physiciens admettaient que l'univers était composé de corpuscules microscopiques et que tous les phénomènes naturels pouvaient être expliqués au moyen des concepts de forme, taille, mouvement et interaction corpusculaires. Cet ensemble d'engagements se révéla être à la fois métaphysique et méthodologique. En tant qu'engagements métaphysiques, ils disaient aux scientifiques quelles sortes d'entités l'univers contenait et ne contenait pas : il y avait seulement de la matière informée en mouvement. En tant qu'impératifs méthodologiques, ils leurs disaient à quoi les lois ultimes et explications fondamentales doivent ressembler : les lois doivent spécifier le mouvement et l'interaction corpusculaires, et l'explication doit réduire tout phénomène naturel donné à une action corpusculaire obéissant à ces lois. (Kuhn, 1962/1970, pp. 41-42)

Ces deux types d'engagements sont parfois, selon Kuhn, indépendants – plus précisément, l'adoption d'un modèle méthodologique n'implique pas l'engagement métaphysique qui peut l'accompagner. Qui plus est, Kuhn affirme que la nature des engagements individuels en faveur d'un modèle n'est pas toujours un critère d'appartenance à un groupe scientifique, dans la mesure où il n'affecte pas les pratiques de ce groupe. Il arrive même parfois que des dissensions explicites existent au sein d'un groupe caractérisé par le partage de modèles analogiques et de généralisations symboliques :

une théorie de la matière n'est pas la sorte de sujet sur lequel les membres d'une même communauté doivent nécessairement s'accorder. *La nécessité de l'accord*

*dépend de ce que fait la communauté.*¹⁷ La chimie de la première moitié du dix-neuvième siècle en fournit un exemple. Bien que beaucoup des outils fondamentaux de la communauté – les proportions constantes, les proportions multiples, la loi des nombres proportionnels – soient devenus la propriété commune des chimistes à la suite de la théorie atomique de Dalton, il était possible pour les chimistes, après l'établissement de cette théorie, de faire reposer leurs travaux sur ces outils et de refuser, parfois avec véhémence, leur adhésion à l'existence des atomes. (Kuhn, 1970a, p. 180)

Si ces désaccords fondamentaux ne remettent pas forcément en question, selon Kuhn, l'appartenance des individus au même groupe, c'est parce que, bien que « la force des engagements du groupe varie, avec des conséquences non triviales, le long du spectre des modèles heuristiques aux modèles ontologiques, *tous les modèles ont des fonctions similaires.*¹⁸ » (Kuhn, 1970a, p. 184) Cela signifie que, quelle que soit la nature des engagements individuels en faveur de ces modèles, l'acceptation d'un modèle comme impératif méthodologique suffit à déterminer l'appartenance à une communauté, *dans la mesure où elle suffit à imposer l'usage de certains types de représentations.*

Insistons sur un point important. On a vu que l'appartenance à un groupe se caractérise, pour Kuhn, par le fait que les représentations utilisées par les membres de ce groupe ont la même signification. Or, on vient de voir que les pratiques concrètes des individus, et en particulier leur choix d'un certain type de représentation, sont des déterminants plus forts de leur appartenance à un groupe que leurs engagements métaphysiques. Cela semble signifier que l'attitude épistémique d'un individu vis-à-vis des représentations qu'il utilise – le fait qu'il croie que ces représentations sont vraies, ou qu'il les considère comme de simples analogies – n'en détermine pas la *compréhension*. Autrement dit, un agent qui croit que la structure fondamentale du monde est faite d'atomes et un agent qui se contente d'adopter un modèle mécaniste comprennent les équations qu'ils utilisent de la même manière.

Afin d'illustrer cela, reprenons l'exemple de la préférence des physiciens de la fin du dix-neuvième siècle pour les explications mécaniques (voir chapitre 4, section 2.1, et chapitre 5, pages 216 et 228). Comme je l'ai déjà souligné (page 172), cette préférence n'implique pas forcément la défense de l'hypothèse atomiste. Pour autant, elle ne consiste pas seulement dans le choix d'une forme de représentation parmi d'autres : elle s'accompagne, par exemple chez Boltzmann, de l'affirmation selon laquelle une théorie que l'on ne peut pas exprimer sous la forme d'un modèle mécanique est proprement, du moins à son époque, *inintelligible*. Cela signifie que l'on ne peut comprendre cette théorie que si elle nous permet de nous représenter les phénomènes

¹⁷ Je souligne.

¹⁸ Je souligne.

comme des phénomènes mécaniques. Que l'on affirme ou non que les phénomènes sont bien tels que ce modèle nous les présente, on ne peut se les représenter – et comprendre la théorie – que sous cette forme. Pour le dire autrement, on peut rester agnostique, par exemple, sur la question de l'existence de forces, tout en étant incapable de se représenter les phénomènes du mouvement autrement qu'en termes de forces, ou, du moins, tout en préférant parler de forces en raison du caractère intuitif de ce concept (voir, sur ce point, mon examen de la version de Mach, chapitre 3, section 4.1.1).

Ces considérations sont importantes pour mon propos à un double titre. D'abord parce que, dans la mesure où ils semblent suggérer que deux individus peuvent comprendre une représentation de la même manière alors même qu'ils n'ont pas les mêmes croyances, ces propos de Kuhn entrent en contradiction avec certaines de ses thèses relatives à l'incommensurabilité des paradigmes, comme on le verra bientôt. Cela me permettra de montrer que cette contradiction révèle les limites de l'approche de Kuhn en tant qu'elle est globaliste.

Ensuite parce que, en attirant l'attention sur la variété des attitudes épistémiques possibles au sein d'un groupe d'agents qui utilisent les mêmes représentations, ces considérations permettent une analyse fine de la manière dont les engagements théoriques de chacun jouent sur leur utilisation de différentes représentations. Si différentes attitudes épistémiques vis-à-vis d'un modèle n'impliquent pas toujours une compréhension et une utilisation différentes de ce modèle, elles peuvent donner lieu, dans certains cas, à des désaccords relatifs à la pratique et à l'utilisation de certaines représentations, désaccords qui impliquent parfois des enjeux théoriques cruciaux. C'est ce que l'analyse des débats des généticiens des années 1920 sur la construction et l'interprétation des cartes génétiques montrera au chapitre 9.

Les valeurs constituent la quatrième composante principale de la matrice disciplinaire. Elles correspondent aux jugements que les individus portent sur les vertus que l'on doit exiger d'une théorie, comme, par exemple, la simplicité, la cohérence, la plausibilité ; la notion de valeur peut aussi désigner, par exemple, la préférence pour les prédictions quantitatives sur les prédictions qualitatives, ou encore l'évaluation de la marge d'erreur autorisée d'une prédiction (Kuhn, 1962/1970, p. 185).¹⁹ Elles jouent, selon Kuhn, un rôle important dans les choix qu'opèrent les scientifiques au moment où une crise survient au sein d'un paradigme.²⁰ Kuhn remarque que ce sont les déterminants les plus globaux, et qu'elles sont généralement « plus large-

¹⁹Elles correspondent assez bien à ce que j'ai appelé, dans mon analyse des versions de la mécanique, les « idéaux scientifiques » des différents auteurs étudiés (chapitre 3, section 5.1).

²⁰Pour la question de l'importance des valeurs dans le choix théorique, voir (Kuhn, 1977b).

ment partagées par les différentes communautés que les généralisations symboliques ou les modèles, [et] fournissent un sentiment d'appartenance communautaire à tous les praticiens des sciences empiriques. » (Kuhn, 1962/1970, p. 185)²¹

1.1.2 Le maintien d'une unité globale

On vient de voir que la notion de paradigme, dans sa grande hétérogénéité, permet d'éclairer plusieurs aspects de la pratique scientifique qui déterminent la compréhension que les agents ont des représentations qu'ils utilisent. L'analyse des différentes composantes de la matrice disciplinaire autorise de plus à identifier des groupes d'individus à différents niveaux de généralité, sur la base des pratiques qui leur sont communes. L'analyse kuhnienne de la manière dont les agents utilisent et comprennent des représentations particulières permet ainsi de rendre compte de l'existence d'authentiques différences conceptuelles au sein de ce que l'on considère habituellement comme une seule théorie – la mécanique quantique n'a pas la même *signification* pour un chimiste et un spécialiste de physique des solides.

Ces différents avantages de l'analyse kuhnienne sont cependant dus à l'examen d'unités locales comme les cas exemplaires et les modèles²², qui permettent comme on l'a vu d'identifier des communautés scientifiques à différents niveaux de généralité. Tout en reconnaissant la polysémie du terme « paradigme » et en affirmant l'importance d'unités locales comme les cas exemplaires, Kuhn maintient pourtant, y compris dans les années 1970, l'idée d'une unité *globale* – la matrice disciplinaire –, qui caractérise l'ensemble des croyances d'une communauté. Comme je l'ai souligné précédemment (page 287), le maintien de cette unité globale semble suggérer qu'elle consiste en « quelque chose de plus » que l'ensemble de ses composantes, ou, plus précisément, que ses composantes, prises ensemble, constituent une unité qui, en tant que telle, caractérise une communauté scientifique. Autrement dit, par delà les différents critères d'identification d'une communauté et par delà l'analyse des pratiques de ses membres, Kuhn affirme l'existence d'une entité conceptuelle globale qui détermine

²¹ Ainsi, on a vu, au chapitre 3 (voir en particulier page 152), que les physiciens philosophes de la fin du dix-neuvième siècle partagent tous un idéal de simplicité et d'économie de la pensée, ainsi qu'un refus des explications « métaphysiques ». Toutefois, cet accord explicite recèle souvent un désaccord sur la manière dont cet idéal doit être mis en œuvre. Si l'on considère, par exemple, les divergences entre les versions de la mécanique exprimées par Duhem et par Mach, qui s'accordent pourtant explicitement sur certaines valeurs, on trouve qu'elles reposent sur des dissensions plus profondes concernant la nature même des théories scientifiques et sur des engagements théoriques différents.

²² Les modèles sont décrits comme des unités globales au sens où ils caractérisent parfois un ensemble de croyances implicites. Cependant, comme on l'a vu, leur fonction au sein d'une communauté consiste principalement à imposer un certain type de généralisation symbolique.

la manière dont les membres de cette communauté comprennent les théories qu'ils utilisent et, par là même, comprennent les phénomènes.

Ainsi défini, le paradigme n'est pas tant une *unité d'analyse* que l'objet d'une *hypothèse théorique*. Or, l'utilité d'une telle hypothèse semble douteuse, puisque l'essentiel des bénéfices de l'analyse qui précède ont été tirés de l'attention à des pratiques et des unités locales. De plus, outre le fait que rien ne permet d'en affirmer l'existence, son maintien – en tant qu'unité globale distincte de l'ensemble de ses composantes – conduit Kuhn, comme on va le voir dans les deux prochaines sections (1.2 et 1.3), à définir le paradigme comme une entité monolithique, restreignant du même coup la fécondité des outils d'analyse que je viens de présenter.

1.2 La thèse de l'incommensurabilité des paradigmes et les limites de l'approche globaliste de Kuhn

On a vu (section 1.1.1) que l'analyse de l'application des généralisations symboliques à des problèmes concrets au cours de l'apprentissage et de la pratique scientifiques permet à Kuhn d'avancer une thèse sur la *signification* des représentations utilisées par les agents. Cette thèse, selon laquelle une expression comme « $f = ma$ » tire sa signification de son application à des situations concrètes que l'étudiant apprend à identifier comme des situations newtoniennes, est parfois mise au service, par Kuhn, d'une thèse générale sur la nature du changement scientifique et des relations inter-théoriques, qui se présente comme une attaque frontale de la conception empiriste logique des théories.

Dans ce cadre, les paradigmes sont identifiés à des schèmes conceptuels qui déterminent à la fois la signification des termes du langage employé par les membres d'une communauté scientifique et le contenu des théories qu'ils formulent au moyen de ce langage ; les paradigmes, qui se confondent alors avec les théories et les langages utilisés pour les exprimer, sont dits « incommensurables », ce qui signifie qu'il n'existe pas de langage commun permettant de les exprimer et, par là, de comparer leur signification.²³

Dans cette section, je vais montrer que cette thèse non seulement ne découle pas, comme Kuhn le prétend, de son analyse de l'apprentissage et de la pratique scientifiques, mais encore repose sur une conception de la signification des théories

²³La thèse de l'incommensurabilité des paradigmes affirme plus largement qu'il n'existe pas d'étalon commun à l'aune duquel des traditions scientifiques successives pourraient être comparées : les normes de rationalité, les méthodes, les valeurs et les croyances fondamentales sont bouleversées. Je me contenterai ici d'examiner la version linguistique de la thèse de l'incommensurabilité, qui fonde les autres aspects de cette thèse et est directement orientée contre la conception des relations inter-théoriques défendue par les empiristes logiques.

différente de celle que j'ai présentée dans ce qui précède. On verra que les efforts que Kuhn consacre à la défense de la thèse de l'incommensurabilité l'écartent de son objet premier qu'est l'étude de la pratique scientifique et le rapprochent paradoxalement de l'empirisme logique dont il prétend pourtant, plus que jamais, se démarquer. Critiquer la thèse de l'incommensurabilité des paradigmes me permettra de montrer les limites de l'approche kuhnienne *en tant qu'elle est globaliste* et qu'elle repose, du moins quand il s'agit d'étudier le changement scientifique et les relations inter-théoriques, sur une conception logico-empirique des théories. Cela me conduira à recentrer mon analyse sur des unités locales et à renoncer à en étudier la signification sur le modèle de la référence linguistique.

1.2.1 De l'application des généralisations symboliques à l'adoption d'un schème conceptuel

Comment s'opère ce glissement qui conduit Kuhn, de ses considérations sur la pratique scientifique, à adopter une conception logico-empirique de la signification des théories ? On a vu, dans ce qui précède, qu'apprendre à appliquer l'expression « $f = ma$ » à des problèmes particuliers, c'est, selon Kuhn, apprendre à reconnaître dans la nature des situations newtoniennes. À cela, Kuhn ajoute parfois deux éléments qui n'en découlent pas nécessairement, et qui reposent sur l'adoption de deux présupposés supplémentaires. Le premier de ces présupposés est l'idée selon laquelle l'apprentissage de la signification d'un énoncé est indissociable de l'acquisition d'une connaissance factuelle – ou d'une croyance – à propos du monde. Le second consiste en l'adoption d'une thèse holiste selon laquelle la signification d'un énoncé lui est fournie par la théorie – ou le système entier de croyances – dont il fait partie.²⁴

À plusieurs reprises, Kuhn affirme que, par l'application de l'expression « $f = ma$ » à des situations concrètes, on apprend, dans le même mouvement, *la signification* des concepts « masse », « force » et « accélération » et *le fait* que le monde est newtonien, c'est-à-dire que « $f = ma$ » exprime une proposition vraie. En effet, selon lui,

Contempler des exemples est une partie essentielle (quoique pas la seule) de l'apprentissage de ce que les mots d'une loi signifient individuellement et collectivement, ou de l'apprentissage de la manière dont il convient de les relier

²⁴Ces deux présupposés correspondent aux thèses au moyen desquelles Quine (1951) dénonce les « deux dogmes de l'empirisme » : il affirme en effet que la distinction entre les questions de fait et les questions de signification est relative au schème conceptuel, et, corrélativement, que la signification des termes et des énoncés est déterminée par le schème conceptuel. Kuhn se réclame explicitement de Quine dans la préface (p. vi) de *La Structure des révolutions scientifiques*. Il y mentionne également l'influence des travaux de Benjamin Whorf (1956) sur le relativisme linguistique.

à la nature. De même, c'est une partie de l'apprentissage *du comportement du monde*²⁵. Les deux ne peuvent pas être séparés. Ce même double rôle est joué par les problèmes des manuels desquels les étudiants apprennent, par exemple, à découvrir forces, masses, accélérations dans la nature et dans le même mouvement découvrent ce que « $f = ma$ » signifie et comment il est attaché à et *gouverne la nature*²⁶. (Kuhn, 1970b, 171)

Quand ces considérations, qui portent sur des formules particulières, sont étendues aux *théories* – dont on a pourtant vu qu'elles n'étaient censées être, tout au plus, qu'un élément parmi d'autres du paradigme – on en arrive à la thèse de l'incommensurabilité des paradigmes.

En effet, à plusieurs reprises, Kuhn affirme que « résoudre des problèmes revient à apprendre *le langage d'une théorie*²⁷ et à acquérir *la connaissance de la nature contenue dans ce langage*²⁸. » (Kuhn, 1970b, p. 169) Dans la mesure où l'apprentissage de la signification des énoncés, se faisant par un « processus analogue à l'ostension », implique d'emblée l'acquisition d'un certain contenu de connaissance sur le monde, il devient impossible de distinguer la théorie (le contenu de connaissance) du langage dans lequel elle est exprimée. En conséquence, une théorie scientifique ne consiste pas en l'expression d'un certain contenu empirique par le biais d'un langage dans lequel il est possible d'exprimer une autre théorie, éventuellement incompatible ou contradictoire avec la première : la théorie scientifique est, dans la perspective de ces propos de Kuhn, *identifiée à un langage*, dont la maîtrise implique d'emblée l'adoption d'un certain nombre de croyances. Le paradigme, en tant qu'il désigne l'ensemble des croyances partagées par une communauté, n'est plus distingué de la théorie, et la communauté scientifique est identifiée à une communauté linguistique.

1.2.2 Incommensurabilité et réduction inter-théorique

La notion de paradigme sert à la fois à décrire ce que les membres d'une communauté scientifique ont en commun et ce qui distingue les communautés – contemporaines ou successives – les unes des autres. Les révolutions scientifiques, c'est-à-dire les changements qui n'ont pas lieu *au sein* de la science normale, mais qui bouleversent cette dernière au point de donner naissance à une nouvelle tradition scientifique, sont donc décrites comme des changements de paradigme.

Dans la mesure où l'appartenance à une communauté scientifique consiste dans la maîtrise d'un langage, laquelle s'accompagne *ipso facto* de l'adoption de croyances

²⁵ Je souligne

²⁶ Je souligne.

²⁷ Je souligne.

²⁸ Je souligne.

et d'une manière de catégoriser et, littéralement, de *voir* le monde, des individus appartenant à des traditions scientifiques différentes, en attribuant des significations différentes aux mots qu'ils emploient, non seulement ne disent pas la même chose du monde, mais ne parlent même pas du même monde. Parce que « ni les scientifiques ni les profanes n'apprennent à voir le monde morceau par morceau ou objet par objet »²⁹, la modification d'une croyance, dans la mesure où elle implique la redéfinition de certains termes, est d'emblée une modification de la manière globale de catégoriser les phénomènes, un « déplacement du réseau conceptuel par lequel le scientifique voit le monde » (Kuhn, 1962/1970, p. 102) :

les coperniciens qui refusaient au Soleil le titre traditionnel de « planète » n'apprenaient pas seulement ce que le terme « planète » signifie ou ce qu'est le Soleil. En fait, ils changeaient la signification même de « planète » de manière à ce que ce terme puisse continuer à établir des distinctions utiles dans un monde où tous les corps célestes, et pas seulement le Soleil, étaient vus sous un aspect différent de celui sous lequel on les avait vus jusque là. [...] Les paradigmes déterminent en même temps de grandes parties de l'expérience. (Kuhn, 1962/1970, pp. 128-129)

C'est cette version linguistique de la thèse de l'incommensurabilité que Kuhn oppose à la conception empiriste logique des relations entre théories, et en particulier à la notion de réduction inter-théorique (voir Nagel, 1961, chap. 11). Cette dernière notion permet d'étudier les relations entre des théories successives ou entre des théories contemporaines portant sur des domaines différents de phénomènes (physique et biologie, par exemple). Elle repose, comme on va le voir, sur l'idée qu'il est possible d'exprimer deux théories différentes dans le même langage, ce qui permet d'en étudier les relations et, le cas échéant, de montrer que l'une est la conséquence logique de l'autre.

Dire qu'une théorie T_1 en réduit une autre T_2 , c'est dire que tous les phénomènes expliqués et prédits par T_2 sont prédictibles et explicables par T_1 , selon le modèle déductif-nomologique de Hempel (1965a).³⁰ Autrement dit, T_2 , pour être réduite par T_1 , doit en être une conséquence logique. Nagel (1961) distingue deux types de réductions ; le premier type, jugé non problématique, est celui des réductions dites « homogènes », c'est-à-dire des cas où tous les concepts de T_2 (la théorie réduite) se

²⁹ « L'enfant qui transfère le mot "maman" de tous les humains à toutes les femmes et finalement à sa mère n'apprend pas seulement ce que "maman" signifie ou qui est sa mère. Simultanément, il apprend certaines différences entre hommes et femmes ainsi que certaines choses sur la manière dont toutes les femmes, sauf une, se comportent vis-à-vis de lui. Ses réactions, ses attentes et ses croyances – en fait, une grande partie du monde qu'il perçoit – changent en conséquence. » (Kuhn, 1962/1970, p. 128)

³⁰ Voir chapitre 5, page 214.

retrouvent dans T_1 .³¹ Le cas du passage de la mécanique newtonienne à la théorie de la relativité d'Einstein est un cas typique, pour Nagel, de réduction homogène.

Dans le chapitre sur « la nature et la nécessité des révolutions scientifiques », Kuhn (1962/1970) propose une critique en règle de la thèse selon laquelle les lois de Newton sont une conséquence logique de celles d'Einstein.³² Selon lui, la position des empiristes logiques est à la fois historiquement improbable et logiquement fautive.³³

Sur le premier point, que je ne développerai pas, l'argument de Kuhn (1962/1970, pp. 98-100) consiste à montrer que la théorie positiviste de la réduction rend proprement impensable le changement scientifique. En définissant la scientificité par la vérifiabilité empirique, les positivistes sont conduits, selon la lecture caricaturale qu'en fait ici Kuhn, à restreindre ce que disent les théories à ce qui a déjà été effectivement vérifié. Ces dernières sont donc par définition prémunies contre l'erreur. Cela rend impossible, selon Kuhn, la mise en défaut de « n'importe quelle théorie ayant été appliquée avec succès à n'importe quel ensemble de phénomènes » (p. 100).

Outre cette incapacité à penser la possibilité même du changement théorique, la position réductionniste souffre, selon Kuhn, d'une « lacune logique » (p. 101). Contrairement à ce que la théorie de la réduction homogène affirme, on ne peut pas déduire la loi de Newton de la théorie d'Einstein, *même sous la forme d'une approximation*.

Moyennant une approximation, on peut certes déduire de la théorie einsteinienne des lois qui ont les mêmes expressions symboliques que celles de Newton. Cependant,

³¹Pour Nagel, les seuls cas problématiques sont ceux des réductions non homogènes, comme celle des lois de la thermodynamique par la théorie cinétique des gaz ; dans ce cas, où le concept de température est « traduit » ou « réduit » par celui d'énergie cinétique, se pose un problème analogue à celui de la réduction des termes théoriques à des termes observationnels (voir chapitre 5, section 1.3) : il semble qu'une partie de la signification de « température » échappe à l'expression « énergie cinétique moyenne des molécules ». Le cas de la « réduction » des lois de Newton par la mécanique relativiste ne pose pas ce problème : la définition des concepts est modifiée, mais, selon Nagel, aucune partie de la signification du concept newtonien de masse n'est perdue dans le processus de réduction.

³²Il considère même cet exemple comme le « prototype des réorientations révolutionnaires dans les sciences » : tout en étant « plus subtile que les changements du géocentrisme à l'héliocentrisme, du phlogistique à l'oxygène, des particules aux ondes, la transformation conceptuelle qui résulte de cette révolution n'est pas moins décisive de la destruction du paradigme établi auparavant. [...] Précisément parce qu'elle n'implique pas l'introduction d'objets ou de concepts additionnels, la transition de la mécanique newtonienne à la mécanique einsteinienne illustre avec une clarté particulière la révolution scientifique comme un déplacement du réseau conceptuel par lequel le scientifique voit le monde. » (Kuhn, 1962/1970, p. 102)

³³Kuhn, comme il le reconnaîtra plus tard, a une connaissance relativement mauvaise des thèses des empiristes logiques, dont il livre une caricature assez grossière. Que la cible de Kuhn soit un homme de paille importe cependant assez peu. Ce qui m'intéresse ici est la conception des théories et de leur signification qui sous-tend son argumentation.

ces lois ne sont pas, selon Kuhn, les lois de Newton. « Tout au plus, ce sont les lois de Newton réinterprétées d'une manière qui était impossible avant les travaux d'Einstein. » (p. 101) En effet, nous dit Kuhn, tout en s'écrivant, par exemple, « $f = ma$ », une expression déduite des lois d'Einstein n'est pas la loi de Newton parce que

les référents physiques [des concepts einsteiniens représentés par les symboles « f », « m » et « a »] ne sont en aucun cas identiques à ceux des concepts newtoniens qui portent le même nom. (La masse newtonienne est conservée; la masse einsteinienne est convertible en énergie. C'est seulement à des vitesses relatives basses que les deux peuvent être mesurées de la même manière, et même dans ce cas elles ne doivent pas être conçues de la même manière.) [...] En passant à la limite, ce ne sont pas seulement les formes des lois qui ont changé. Simultanément, nous avons dû altérer les éléments structuraux fondamentaux dont l'univers auquel elles sont appliquées est composé. » (Kuhn, 1962/1970, pp. 101-102)

Du fait que le concept de masse n'a pas la même définition dans la théorie de Newton et dans celle d'Einstein, et donc que l'expression « $f = ma$ » n'a pas la même signification dans les deux contextes, Kuhn déduit l'affirmation selon laquelle le symbole « m » n'a pas le même référent. Cet argument de Kuhn repose sur les deux présupposés mentionnés précédemment : d'une part, l'indistinction entre les questions de signification et les questions de fait, et d'autre part l'adoption du holisme sémantique. À en croire Kuhn, puisque changement de signification des termes et changement de croyance sur le monde sont indissociables, la signification du concept de masse chez Einstein ne peut pas être comprise si l'on n'adopte pas la théorie d'Einstein et la vision du monde qui l'accompagne.

Adopter la théorie de Newton ou la théorie d'Einstein, c'est adopter d'emblée un ensemble de croyances et une vision globale du monde dans laquelle les termes de l'autre théorie n'ont pas de référent.

Dans la transition d'une théorie à une autre, les mots changent de signification ou de conditions d'applicabilité d'une manière subtile. Bien que la plupart des signes utilisés avant et après une révolution soient les mêmes – par exemple force, masse, élément, composé, cellule – les manières selon lesquelles certains d'entre eux sont reliés à la nature ont quelque peu changé. Les théories successives sont donc, comme nous [Feyerabend et Kuhn]³⁴ le disons, incommensurables. (Kuhn, 1970b, p. 163)

³⁴Dans ce texte, Kuhn n'opère pas de distinction nette entre la thèse de l'incommensurabilité défendue par Feyerabend (1962) et la sienne. Dans d'autres textes, comme par exemple dans (Kuhn, 1982), il s'en démarque plus nettement, signalant que Feyerabend défend la thèse radicale selon laquelle la signification de *tous* les termes, y compris ceux qui semblent désigner des observables de manière indépendante de la théorie, est affectée par un changement de théorie. Kuhn, pour sa part, prétendra limiter l'incommensurabilité à certains termes seulement. Il est cependant difficile, dans

La thèse selon laquelle le donné est toujours d'emblée théorisé, quand elle prend la forme d'un holisme sémantique, conduit donc Kuhn à l'affirmation radicale selon laquelle mécaniques newtonienne et einsteinienne non seulement ne disent pas la même chose, mais encore ne parlent pas du même monde ; tout comme l'astronomie de Ptolémée et celle de Copernic, les théories de Newton et d'Einstein sont selon lui fondamentalement incompatibles.

1.2.3 Retour à une conception linguistique des théories

Je voudrais à présent montrer que cet argument repose sur une concession majeure à la conception empiriste logique des théories. En effet, si le terme « masse » n'a pas la même signification chez Newton et chez Einstein, c'est entre autres, nous dit Kuhn, parce que « la masse newtonienne est conservée » alors que « la masse einsteinienne est convertible en énergie ». La signification de ce terme lui est donc donnée par la théorie qui l'emploie, cette théorie étant conçue comme un ensemble d'énoncés qui tout à la fois fixe les significations des mots et exprime des propositions sur le monde. La concentration exclusive de Kuhn sur les *théories* et le fait que la thèse de l'incommensurabilité s'applique à elles et non à l'ensemble, dont on a vu qu'il était extrêmement hétérogène, des composantes du paradigme, est le symptôme d'une dérive linguistique des propos kuhnien, qui leur fait perdre, selon moi, une partie de leur force.

Ces propos semblent en effet incompatibles avec une autre conception de la signification des lois et des concepts scientifiques, selon laquelle la signification de « $f = ma$ », pour l'individu qui l'utilise, dépend de la manière dont il a appris à l'appliquer à des situations concrètes. Kuhn insiste en effet, on l'a vu, sur le fait que l'étudiant en physique « découvre la signification de termes comme “force”, “masse”, “espace” et “temps” [...] moins par le biais des définitions incomplètes quoique parfois utiles qu'il trouve dans son manuel qu'en observant et en participant à l'application de ces concepts à des solutions de problèmes » (Kuhn, 1962/1970, p. 47). Dès lors, on voit mal pourquoi les définitions incompatibles de la masse par les théories newtonienne et einsteinienne impliquent que, même dans le cas de « vitesses relatives basses [où les masses newtonienne et einsteinienne] *peuvent être mesurées de la même manière*³⁵, [...] elles ne doivent pas être conçues de la même manière. » Si vraiment « $f = ma$ » est une expression vide de signification avant tout exercice concret de résolution de problèmes, on devrait pouvoir, suivant Kuhn, affirmer que, quelle que

l'argument qu'il développe en 1962 contre la réduction inter-théorique, de voir comment une telle limitation peut être imposée au changement de signification.

³⁵ Je souligne.

soit la théorie d'où elle est issue, sa signification est la même quand son application implique les mêmes manipulations de la part du scientifique. Pourtant, dans l'argument de Kuhn, si les théories newtonienne et einsteinienne sont incommensurables, ce n'est pas parce que l'expression « $f = ma$ » n'est pas appliquée aux mêmes cas et n'implique pas les mêmes manipulations dans le cadre newtonien et dans le cadre einsteinien, mais parce que les différentes lois dans lesquelles figure le symbole « m » ne sont pas *logiquement* équivalentes.

Alors même que ses propos sur le rôle des généralisations symboliques et les cas exemplaires étaient destinés à rejeter la pertinence de la notion de théorie entendue comme un système déductif d'énoncés, Kuhn semble ici opérer un retour à cette conception des théories. Le holisme sémantique qui sous-tend la thèse de l'incommensurabilité, en permettant d'affirmer que, quand la signification d'un seul terme est modifiée, le système entier ainsi que la structure du monde qu'il décrit sont bouleversés, repose en fait, comme on l'a vu, sur l'identification préalable d'une théorie à un langage. La critique de Kuhn aux empiristes logiques se « réduit » ici à celle de Quine (1951), qui affirme l'impossibilité d'une distinction entre questions de fait et questions de signification. Sans minimiser l'importance d'une telle critique, je suggère qu'elle n'a de sens qu'en tant qu'elle repose sur l'acceptation d'une part importante de l'héritage des empiristes logiques. Elle repose en particulier sur une approche exclusivement linguistique de la connaissance scientifique, conduisant à des questions de philosophie de la logique et du langage qui n'ont qu'une incidence lointaine sur la réalité de la pratique scientifique.

En effet, la thèse de l'incommensurabilité ne contribue pas à donner une image plus réaliste de la pratique scientifique que celle des empiristes logiques orthodoxes. Si l'idée d'un ensemble de règles de correspondance fournissant une interprétation empirique à un système formel ne rend pas compte de la nature de l'apprentissage et de la connaissance scientifiques, celle d'un schème conceptuel qui implique de la part des individus l'adoption globale d'un ensemble de croyances n'est pas plus satisfaisante. Contrairement à ce que prétend Kuhn, elle ne rend pas compte de la manière dont la connaissance scientifique est construite et acquise, « morceau par morceau » : l'affirmation selon laquelle la signification de « masse » n'est pas la même chez Newton et chez Einstein n'est en aucun cas une affirmation sur la manière dont les énoncés dans lesquels ce terme figure sont utilisés et compris par les agents, mais bien sur la différence logique entre deux systèmes formels.

1.2.4 Le changement scientifique rendu incompréhensible

Rappelons la double critique que Kuhn faisait à la conception empiriste logique de la réduction inter-théorique (voir page 301) : elle est historiquement improbable et souffre d'une lacune logique. Du point de vue logique, la critique kuhnienne repose, on vient de le voir, sur l'adoption des présupposés fondamentaux de ses adversaires. Je voudrais à présent montrer que, du point de vue historique, il en découle une conception tout aussi irréaliste.

Affirmer que les théories de Newton et d'Einstein, parce qu'elles ne disent pas la même chose, ne parlent tout simplement pas de la même chose revient à nier qu'il y ait progrès ou même changement de Newton à Einstein. À la rigueur, rien ne permet de dire qu'elles ont plus en commun que ne l'ont, par exemple, l'astronomie et la botanique. Si « masse », chez Newton et chez Einstein, ne désigne pas la même réalité, on voit mal en quoi consiste le *changement*.

Dudley Shapere (1964, 1966) a ainsi montré que la thèse de l'incommensurabilité revient à faire du changement scientifique, assimilé à un changement de signification, une affaire de « tout ou rien ». Cela, joint à l'absence de critère clair de changement de signification³⁶, contribue selon Schapere à faire de la notion de paradigme une unité d'analyse inopérante. La « racine du relativisme de Kuhn et de Feyerabend », à laquelle la thèse de l'incommensurabilité radicale conduit, réside selon Shapere (1966) dans leur « conception rigide de ce en quoi consiste un changement de signification – à savoir, incomparabilité absolue, “incommensurabilité”. » (pp. 67-68)

Deux expressions ou ensemble d'expressions doivent soit avoir exactement la même signification, soit être radicalement et complètement différent(e)s. Si la signification des théories n'est pas invariante au cours de l'histoire de leur développement et de leur incorporation dans des théories plus larges ou plus pro-

³⁶ Le holisme sémantique ne fournit en effet, selon Schapere, aucun critère permettant de distinguer un changement dans la signification d'un terme d'un changement dans son application, ce qui rend arbitraire tout jugement affirmant un changement révolutionnaire. « On ne nous fournit aucun moyen de décider de ce qui compte comme une partie de la “signification” d'un terme ou de ce qui compte comme un “changement de signification” de ce terme. De même, on ne nous fournit aucun moyen de décider de ce qui compte comme une partie d'une “théorie” et de ce qui compte comme un “changement de théorie” » (Shapere, 1966, p. 55). Cette critique est adressée, dans ce contexte, à Feyerabend, dont la thèse de l'incommensurabilité est, plus nettement que chez Kuhn, orientée contre la notion nagélienne de réduction inter-théorique. Dans la mesure où Kuhn, dans d'autres textes, distingue sa thèse de l'incommensurabilité de celle de Feyerabend (voir note 34), la critique de Schapere peut sembler injuste quand elle s'adresse à Kuhn. Cependant, dans les textes où Kuhn, de façon symptomatique, se concentre sur les théories et non sur les paradigmes, sa thèse est parfois difficile à distinguer de celle de Feyerabend. Ce dernier, choisissant d'appeler « théorie » son unité d'analyse globale, énonce sa thèse sur un plan uniquement linguistique et en assume entièrement les conséquences irrationnelles et relativistes, ce qui n'est pas le cas de Kuhn.

fondes, alors ces théories (paradigmes) successifs ne peuvent *réellement* pas être comparés du tout, malgré les apparentes similitudes, qui doivent être négligées car considérées non pertinentes et superficielles. Si la conception de l'histoire des sciences comme un « développement par accumulation » est incorrecte, la seule alternative est qu'il s'agisse d'un processus absolument non cumulatif de remplacement. Il n'y a jamais de demie-mesure et, en conséquence, il n'est pas surprenant que le rejet des principes positivistes d'invariance de la signification et de développement par accumulation nous abandonne au relativisme, car c'est la seule possibilité que cette conception de la différence de signification laisse ouverte. Mais ce relativisme, et les doctrines qui y conduisent, ne sont pas le résultat d'une enquête sur la science réelle et son histoire; au contraire, c'est la conséquence purement logique d'une préconception étroite de la « signification ».

(Shapere, 1966, p. 68)

Selon cette critique – que j'endosse –, Kuhn pêche précisément par là où il accuse les empiristes logiques de pécher. En se concentrant exclusivement sur des questions de signification, et en adoptant une conception très rigide de la signification – qu'il n'explicite pas vraiment –, il se rend incapable de rendre compte de la dynamique de la science. Les théories newtonienne et einsteinienne sont des entités tout aussi figées et abstraites dans sa perspective que dans celle du positivisme logique. À son tour, Kuhn fait des exemples historiques des « anecdotes »³⁷ destinées à illustrer des théories préconçues.

1.3 Paradigmes et versions : deux conceptions de la signification

On a vu (section 1.1) que l'analyse kuhnienne de la pratique scientifique permet de tracer différents types de distinctions au sein de l'activité théorique, en fonction des types d'engagements et de pratiques auxquels on prête attention. J'ai notamment montré qu'elle permet de rendre compte de différences conceptuelles au sein d'une même « théorie », différences gommées par l'analyse formelle des empiristes logiques qui identifient les différences conceptuelles à des différences logico-empiriques. Pourtant, on vient de voir que la conception de la signification qui sous-tend la thèse de l'incommensurabilité conduit Kuhn à dresser un tableau de l'activité scientifique aussi irréaliste que celui des empiristes logiques, en assimilant, à son tour, changement conceptuel et changement logico-empirique. Dans cette section, je me propose d'approfondir cette critique en montrant, dans un premier temps (section 1.3.1), que la thèse de l'incommensurabilité, en faisant des paradigmes des entités monolithiques,

³⁷ Kuhn (1962/1970, p. 1) écrit, dans l'introduction de son ouvrage : « L'histoire, si on la considérait comme autre chose qu'un ensemble d'anecdotes ou de dates, pourrait provoquer une transformation décisive dans l'image de la science dont nous sommes actuellement prisonniers. »

empêche non seulement de comprendre les changements de théorie (comme le passage de la mécanique newtonienne à la mécanique einsteinienne), mais aussi la variabilité conceptuelle *intra-théorique*, que mon analyse des versions de la mécanique, au chapitre 3, a mise en évidence. Dans un deuxième temps, je reviendrai sur les différentes conceptions de la signification que Kuhn, dans ses propos, semble adopter tour à tour, et je les comparerai à celle qui me semble la plus adaptée à la description de l'activité théorique (section 1.3.2). Enfin, je préciserai la comparaison entre les notions de version et de paradigme et les hypothèses qu'elles permettent d'énoncer à propos de l'activité théorique (section 1.3.3).

1.3.1 Les paradigmes comme unités monolithiques

De même qu'ils conduisent à nier toute similitude entre deux théories jugées incommensurables comme celles de Newton et d'Einstein, les propos de Kuhn que j'ai présentés à la section 1.2 empêchent, comme l'a montré Dudley Shapere (1964, 1966), d'apercevoir des différences conceptuelles au sein d'une même tradition scientifique. Ainsi effacent-ils la diversité de ce que j'ai appelé, au chapitre 3, les « versions » – successives ou simultanées – de la mécanique :

il est naturel et courant de dire que Newton, d'Alembert, Lagrange, Hertz, Hamilton, Mach et d'autres ont formulé des versions³⁸ différentes de la mécanique classique ; cependant, certaines de ces formulations impliquaient indiscutablement des « engagements » différents – par exemple, certaines impliquaient des engagements pour les forces, d'autres pour l'énergie, certaines pour des principes vectoriels, d'autres pour des principes variationnels. La distinction entre paradigmes et différentes articulations d'un paradigme, et entre révolutions scientifiques et science normale, est au mieux affaire de degré, tout comme l'engagement pour un paradigme : l'expression de désaccord explicite, la prolifération d'articulations concurrentes, les débats sur les fondements, sont tous plus ou moins présents tout au long du développement de la science ; et il y a toujours des éléments directeurs qui sont plus ou moins communs, même entre ce que l'on considère comme des « traditions » différentes. (Shapere, 1964, pp. 387-388)

En effet, l'idée selon laquelle l'adoption d'un paradigme implique d'emblée celle d'un ensemble de croyances conçu comme une « vision du monde » globale, masque les variations inter-individuelles mises en évidence par l'analyse des versions. Il n'est plus alors possible, au sein d'une communauté, de dégager des sous-communautés ni, au sein d'une « même » théorie, d'identifier des changements conceptuels. On ne

³⁸Dudley Shapere emploie ici indifféremment les termes « versions » et « formulations », et je ne prétends pas lui attribuer les définitions que j'ai proposées dans la première partie de ce travail. Pour autant, les différences qu'il souligne sont proches de celles que j'ai pu établir au cours de mon analyse.

peut qu'identifier de grands paradigmes incommensurables, au sein desquels règne un accord qui rend toute discussion impossible. Puisque deux personnes exprimant des opinions différentes ne parlent d'emblée pas de la même chose, le dialogue est impossible, à l'intérieur comme à l'extérieur du paradigme.

Si Kuhn manque les différences intra-théoriques pourtant caractéristiques de l'activité scientifique normale, c'est, selon Shapere (1966), parce que, comme on l'a vu, il fait du changement scientifique une affaire de « tout ou rien », et parce qu'il suppose l'existence, au fondement de l'activité scientifique d'une communauté d'individus, d'une « unité cachée » et monolithique (comme je l'ai souligné page 287) :

Une fois encore [...], Kuhn commet l'erreur de croire qu'il n'y a que deux branches à l'alternative : identité absolue ou différence absolue. Mais les données disponibles sont les similitudes et les différences ; et pourquoi ne pourraient-elles pas suffire à nous permettre de parler de conceptions plus ou moins similaires, et, selon le but recherché, à classer des points de vue suffisamment similaires ensemble, comme, par exemple, appartenant à la même tradition ? [...] En rigidifiant la notion de « tradition scientifique » en une unité cachée, Kuhn est forcé, *pour une raison purement conceptuelle* à ignorer beaucoup de différences importantes entre activités scientifiques considérées comme appartenant à la même tradition, ainsi que des continuités importantes entre traditions successives. (Shapere, 1966, p. 71)

Kuhn aperçoit pourtant, dès 1962, le risque que la rigidification de la notion de paradigme fait courir à une description fidèle de l'activité scientifique. Il se défend explicitement de proposer une vision monolithique des traditions scientifiques et souligne le fait que la notion de paradigme permet de mieux rendre compte de la complexité de la pratique scientifique que l'idée selon laquelle l'application des théories s'apparente à la mise en œuvre d'un ensemble de règles.

Si la science normale est aussi rigide et si les communautés scientifiques sont si unies que la discussion précédente le suggère, comment un changement de paradigme pourrait-il affecter seulement un sous-groupe ? Ce qui a été dit jusqu'ici peut sembler suggérer que la science normale est une entreprise monolithique unifiée et isolée qui doit tenir ou tomber avec chacun de ses paradigmes, et avec tous en même temps. Mais la science n'est que rarement voire jamais ainsi. Le plus souvent, quand on regarde les différents domaines ensemble, elle ressemble plutôt à une structure branlante dont la cohésion entre les différentes parties est faible. Rien de ce que j'ai dit jusqu'ici ne doit, cependant, entrer en conflit avec cette observation familière. Au contraire, substituer les paradigmes aux règles doit permettre de comprendre plus aisément la diversité des domaines scientifiques et des spécialités. Les règles explicites, quand elles existent, sont généralement communes à un groupe scientifique très large, ce qui n'est pas forcément le cas des paradigmes. [...] Même des hommes appartenant à des do-

maines assez proches et dont l'apprentissage a commencé par la lecture des mêmes livres [...] peuvent acquérir des paradigmes assez différents au cours de leur spécialisation professionnelle.

[...] Une révolution peut se produire au sein d'une de ces traditions sans s'étendre nécessairement aux autres. (pp. 49-50).³⁹

Ces lignes, qui pourraient bien figurer dans une critique de la thèse de l'incommensurabilité, mettent ainsi en évidence la chose suivante : tant que, pour reprendre les termes de Kuhn, le paradigme est décrit comme une « structure branlante », il nous offre le moyen d'opérer des distinctions intéressantes, impossibles à établir dans le cadre empiriste logique. En revanche, quand, au lieu de servir à l'étude des pratiques de la science normale, cette notion est mise au service d'une analyse des révolutions scientifiques, le double mouvement de rigidification et d'identification du paradigme à une entité linguistique globale lui fait perdre l'hétérogénéité et la plasticité qui en faisaient la fécondité.

Comme je l'ai souligné précédemment (section 1.2.3), ces propos contradictoires de Kuhn reposent sur l'adoption de deux conceptions différentes de la signification. Je me propose, dans la section suivante, d'approfondir l'analyse de ce problème, et de comparer ces deux conceptions avec celle qui sous-tend, au chapitre 3 (section 1), mon analyse des différences conceptuelles au sein de la mécanique classique.

1.3.2 Sémantique vériconditionnelle et sémantique inférentielle

Comme on l'a vu (page 291), l'analyse de la manière dont les généralisations symboliques sont appliquées à des cas exemplaires au cours de l'apprentissage et de la pratique scientifiques conduit Kuhn à affirmer que « ce que signifie la mécanique quantique pour chacun [...] dépend de l'enseignement qu'il a suivi, des textes qu'il a lus, et des revues qu'il fréquente » (Kuhn, 1962/1970, p. 50). Une telle conception de la signification semble permettre de dire que les reformulations successives de la mécanique, parce qu'elles ont transformé la manière dont les équations du mouvement peuvent être utilisées (leur « mode d'emploi »), impliquent des changements de signification. Pourtant, Kuhn traite explicitement le cas du développement historique

³⁹Kuhn écrira par ailleurs en 1970 : « En partie en raison des exemples que j'ai choisis et en partie à cause du vague laissé autour de la nature et de la taille des communautés pertinentes, quelques lecteurs de ce livre ont conclu que mon intérêt résidait d'abord, voire exclusivement, dans les révolutions majeures comme celles qu'on associe à Copernic, Newton, Darwin ou Einstein. Une délimitation plus claire de la structure d'une théorie devrait, cependant, aider à renforcer l'impression fort différente que j'ai voulu donner. Une révolution est, pour moi, un type spécial de changement impliquant un certain type de reconstruction des engagements du groupe. Mais elle n'a pas besoin d'être un grand changement, ni de sembler révolutionnaire à ceux qui ne font pas partie d'une certaine communauté, qui peut être composée de vingt-deux personnes. » (Kuhn, 1970a, pp. 180-181)

de la mécanique classique comme un exemple typique du travail d'« articulation » du paradigme dans le cadre de la science normale, qu'il considère comme un travail de développement purement formel :

Ainsi, d'Euler et Lagrange au dix-huitième siècle jusque Hamilton, Jacobi et Hertz au dix-neuvième, plusieurs génies européens de la physique mathématique ont successivement entrepris de reformuler la théorie mécanique *sous une forme logiquement équivalente*⁴⁰ mais esthétiquement plus satisfaisante. Ils souhaitaient, en fait, présenter les leçons explicites et implicites des *Principia* et de la mécanique continentale sous une version logiquement plus cohérente, qui serait à la fois plus uniforme et moins équivoque dans ses applications aux problèmes de mécanique nouvellement élaborés. (Kuhn, 1962/1970, p. 33)

L'équivalence *logique* des formulations de la mécanique en fait, selon Kuhn, des présentations différentes d'un seul et même paradigme, qui n'en modifient pas la signification (de même que la différence *logique* entre les définitions qu'elles donnent de la masse font des théories de Newton et d'Einstein des paradigmes incommensurables).

Ce n'est pourtant, comme je l'ai souligné plusieurs fois, que d'un point de vue abstrait que l'on peut ainsi affirmer l'équivalence des formulations de la mécanique. Si l'on prête attention à la manière dont elles sont reliées aux phénomènes empiriques et dont elles sont utilisées pour résoudre des problèmes, on devrait pouvoir affirmer, suivant les propos de Kuhn sur les cas exemplaires, que les équations de la mécanique n'ont pas la même signification. Du point de vue de l'apprentissage et de la pratique scientifiques, le fait qu'elles soient logiquement équivalentes n'a qu'une pertinence limitée.

Insistons un instant sur le caractère incompatible des deux conceptions de la signification que Kuhn semble adopter tour à tour. Selon l'une d'elle, l'identité de signification est une identité logico-empirique. Dans cette perspective, la notion de force a strictement la même signification, qu'elle s'applique à la force de gravitation expliquant le mouvement d'un corps en chute libre ou aux forces de contrainte maintenant les liaisons d'un système ; en revanche, même si elle est utilisée pour la description du même problème (par exemple le mouvement d'un corps en chute libre), la notion de masse n'a pas du tout la même signification (et n'a pas le même référent) selon qu'elle est utilisée dans le cadre newtonien ou einsteinien. D'ailleurs, dans un cas, on décrit un objet en chute libre, dans l'autre, une déformation riemannienne de l'espace-temps.

Selon l'autre conception de la signification, c'est exactement le contraire : les différentes applications du concept de force contribuent à lui donner une signification, non pas entièrement, mais partiellement différente. En revanche, l'idée d'un « cadre

⁴⁰ Je souligne.

newtonien » ou d'un « cadre einsteinien » dans lequel l'utilisateur de l'expression « $f = ma$ » se situerait pour décrire le mouvement d'un pendule n'a pas vraiment de pertinence pour l'analyse des raisonnements qu'il conduit.

Ainsi, affirmer que la signification du concept de force est *entièrement différente* dans les paradigmes newtonien et einsteinien, et qu'elle est *entièrement identique* dans les formulations newtonienne, lagrangienne et hamiltonienne implique de se situer du point de vue abstrait de la théorie prise comme un tout et non de celui de l'agent qui l'utilise pour décrire les phénomènes. Cette affirmation repose sur la double adoption du holisme sémantique dont on a vu qu'il était au fondement de la thèse de l'incommensurabilité et d'une conception vériconditionnelle ou référentielle de la signification, selon laquelle deux énoncés qui ont les mêmes conditions de vérité « en principe » (et les concepts qui y figurent) ont la même signification.

Signalons que Kuhn (1982), en réponse aux nombreuses critiques⁴¹ qui ont été faites à la notion d'incommensurabilité, situe explicitement la racine de ce qu'il présente alors comme un malentendu dans un certain usage de la notion de signification. Proposant de parler d'« incommensurabilité partielle » (Kuhn, 1982, p. 670) et non totale, il suggère de renoncer à une conception de la signification « extensionnelle, et par conséquent restreinte à la conservation de la valeur de vérité » (p. 680). Kuhn (1982) affirme qu'il préfère éviter « entièrement de parler de “signification” » pour examiner « comment les membres d'une communauté linguistique sélectionnent les référents des termes qu'ils emploient » (p. 681). Cela conduirait, selon lui, à trouver une forme de « holisme local » (p. 682) dans l'usage du langage. Tout en continuant à identifier une communauté scientifique à une communauté linguistique, Kuhn suggère ainsi que la solution aux difficultés auxquelles conduit la thèse de l'incommensurabilité – que Feyerabend, de son côté, assume parfaitement, dans la perspective irrationnaliste et anarchiste qui est la sienne – réside dans une attention à la manière dont, localement, les énoncés des théories sont attachés à la nature ; cela indique clairement un renoncement aux présupposés qui sous-tendent ses propos de 1962 sur l'incommensurabilité des théories newtonienne et einsteinienne.

⁴¹ Voir par exemple (Davidson, 1974; Kitcher, 1978; Putnam, 1981). Hilary Putnam (1973, 1975) développe, à la fois contre les théories de la signification des termes théoriques des empiristes logiques et contre la thèse de l'incommensurabilité, une théorie causale de la référence des termes d'espèces naturelles inspirée de la théorie de la référence des noms propres de Saul Kripke (1980). Philip Kitcher (1982a) propose lui-même une version amendée de la théorie de Putnam, parlant non plus de référence mais de « potentiel de référence », dans la définition duquel le contexte d'énonciation du terme intervient. Il applique notamment cette théorie au problème de la référence du terme « gène ». La critique de Donald Davidson (1974) est adressée à toutes les théories de la signification faisant intervenir la notion de schème conceptuel, qu'il dénonce comme un troisième dogme de l'empirisme.

Je voudrais à présent approfondir cette suggestion en revenant sur l'analyse que j'ai proposé des différences conceptuelles entre les formulations de la mécanique au chapitre 3 (section 1). En insistant sur les différences architectoniques entre ces formulations, j'y ai suggéré que la signification des principes de la mécanique et des concepts qui y figurent dépend non seulement des autres principes de la théorie, mais encore de la place particulière que chacun d'entre eux occupe dans l'ordre déductif de ces différents principes (section 1.1). J'ai alors souligné la différence entre cette conception, que j'ai proposé de qualifier de « holisme *situé* » et le holisme de Kuhn qui ne prête attention qu'aux relations logiques et non aux relations architectoniques (si l'on entend par là les différences d'ordre déductif). J'ai enfin suggéré que la position que je défends implique de renoncer à une conception vériconditionnelle de la signification au profit d'une conception qui accorde de l'importance aux inférences qu'un énoncé permet de tirer en pratique, et non pas seulement à l'ensemble de ses conséquences déductives en principe (section 1.3).

Quelle relation cette conception de la signification entretient-elle avec celle qui permet de dire qu'une généralisation symbolique, pour reprendre le vocabulaire de Kuhn, acquiert sa signification des problèmes concrets auxquels elle est appliquée ? En opposant le holisme de Kuhn à ses considérations sur l'utilisation des cas exemplaires, et en insistant sur la pertinence de celles-ci, je ne prétends pas affirmer que les équations de la mécanique, par exemple, acquièrent leur signification *exclusivement* de leur application à des cas concrets. Il est indéniable qu'une partie de la signification de l'expression « $f = ma$ » lui vient de celle des autres principes auxquels elle est reliée dans le cadre théorique de la mécanique.

Ces deux exigences – celle selon laquelle la signification d'une expression lui vient en partie de son application à des cas concrets et celle selon laquelle elle lui est fournie par sa place dans l'architecture de la théorie – ne sont pas incompatibles. On peut en effet affirmer que la signification d'une expression comme celle de la deuxième loi de Newton lui vient de l'usage que les agents en font dans leurs raisonnements (ces raisonnements étant contraints, rappelons-le, par un certain nombre de relations logiques entre les concepts), qu'il s'agisse de déduire un principe d'un autre dans le cadre du réseau conceptuel de la mécanique ou de l'utiliser pour résoudre un problème concret. Dans les deux cas, cette conception insiste sur les procédures pratiques d'utilisation de cette expression, par opposition à l'idée selon laquelle sa signification lui viendrait exclusivement, d'une part, des relations logiques *en principe* qu'elle entretient avec les autres énoncés de la théorie, et d'autre part de ses conditions de vérité empirique.

Signalons pour finir que cette conception de la signification, qui sous-tend l'ensemble de mon travail, trouve une source d'inspiration dans les propos de plusieurs

philosophes du langage et de l'esprit (entre autres Harman, 1982a, 1987; Greenberg et Harman, 2006; Field, 1977; Block, 1986, 1987; Peacocke, 1992) ainsi que certains chercheurs en psychologie cognitive (par exemple Miller et Johnson-Laird, 1976) qui défendent ce que l'on appelle une « sémantique des rôles conceptuels » ou « sémantique des rôles inférentiels ». Cette perspective, inspirée de l'idée wittgensteinienne (Wittgenstein, 1953) selon laquelle la signification d'un terme est la manière dont il est utilisé, vise à rendre compte du contenu des concepts en des termes autres que vériconditionnels. Bien qu'il s'agisse plus, comme le souligne Ned Block (1998), d'un cadre que d'une théorie, et qu'il existe plusieurs désaccords importants entre ses défenseurs, on peut grossièrement présenter cette conception de la signification comme l'idée selon laquelle le contenu d'un concept est le rôle qu'il joue dans les inférences, et plus largement dans la vie cognitive des agents. La sémantique des rôles conceptuels prend plusieurs formes, plus ou moins radicales, qui se distinguent en particulier par ce qu'elles font entrer dans l'idée large de « vie cognitive » (raisonnements, perception, prise de décision, etc.), par les limites qu'elles assignent aux inférences qui comptent dans la détermination de la signification d'un concept, et par la place qu'elles accordent éventuellement à une composante référentielle.⁴² Dans une version assez libérale comme celle de Greenberg et Harman (2006), la signification d'un concept est le rôle qu'il joue dans les raisonnements des agents au sens large : ce rôle inclut aussi bien la manière dont ce concept est relié au monde que les processus inférentiels dans lesquels il intervient.

Faire de la sémantique des rôles conceptuels un véritable outil pour mon propos impliquerait de prendre en compte les nombreuses objections dont elle a fait l'objet⁴³, ce qui dépasserait le cadre de ce travail. Je souhaite simplement souligner le fait que les motivations initiales de la sémantique des rôles conceptuels, ainsi que plusieurs réflexions de ses défenseurs, offrent des pistes intéressantes pour mon propos.⁴⁴ Ainsi par exemple Ned Block (1998), après avoir signalé que l'une des motivations de la

⁴²Ainsi par exemple la théorie des « deux facteurs » [*two-factors*] considère que la signification d'un concept est constituée d'un aspect interne – le rôle que ce concept joue « à l'intérieur de la tête » des agents – et d'un facteur externe décrit en termes référentiels (voir Block, 1987). On peut aussi, comme Harman (1987), défendre une conception « mono-facteur » [*one factor*] tout en concevant la notion de rôle inférentiel de manière suffisamment large pour que cela permette de s'accommoder de certains problèmes auxquels l'abandon de la sémantique vériconditionnelle conduit.

⁴³Voir par exemple (Fodor et LePore, 1992, 2002).

⁴⁴En outre, plusieurs études en psychologie cognitive et développementale sur l'acquisition des concepts et le changement conceptuel chez les jeunes enfants se présentent comme la mise en œuvre de cette perspective sur la signification et lui offrent une base empirique (voir par exemple Carey, 1985; Gopnik, 1996). Ces études reposent souvent sur une analogie entre le développement conceptuel des individus et le changement conceptuel en science.

sémantique des rôles conceptuels est de rendre compte du fait, souligné par Kuhn, que l'on apprend souvent la signification des concepts scientifiques par leur utilisation dans la résolution de problèmes, suggère que l'on peut s'accommoder de certaines difficultés soulevées par l'abandon d'une sémantique vériconditionnelle en adoptant l'idée d'une « similitude graduelle » de signification plutôt que de n'envisager que l'identité ou la différence strictes. Cela permet en particulier de rendre compte du fait que les locuteurs d'une langue naturelle, tout en utilisant parfois le vocabulaire de différentes manières, s'entendent sur une grande partie de la signification des mots qu'ils emploient. Dans le cas des représentations scientifiques, admettre l'idée d'une similitude graduelle, ou partielle, de signification permet d'éviter les problèmes que Schapere a soulignés dans l'analyse du changement scientifique par Kuhn (voir ci-dessus, pages 307-308), qui donne, comme on l'a vu, une image insatisfaisante de l'activité théorique au sein d'une même tradition ; cela permet de parler, comme le suggère Schapere, de « conceptions plus ou moins similaires », et de rendre compte du fait que la « distinction entre paradigmes et différentes articulations d'un paradigme, et entre révolutions scientifiques et science normale, est au mieux affaire de degré ».

Tout en réhabilitant une forme de holisme « situé » et « local », cette perspective ne conduit pas à renoncer à l'idée selon laquelle deux représentations qui sont, du point de vue global de la théorie, logiquement équivalentes, peuvent ne pas avoir le même contenu pour un individu donné. En outre, comme j'y reviendrai à la fin de ce chapitre et au suivant, l'idée selon laquelle la signification d'un concept est le rôle qu'il joue dans les raisonnements permet des analyses fécondes des représentations non linguistiques.

1.3.3 Paradigmes et versions

Ces considérations permettent de conclure cette section par une comparaison entre les notions de version et de paradigme. J'ai suggéré précédemment (page 287) que la notion de paradigme, en tant qu'elle ne se réduit pas aux différents déterminants de l'appartenance à une communauté scientifique (généralisations symboliques, cas exemplaires, modèles, valeurs), désigne une entité hypothétique. En tant que telle, elle ne peut pas être étudiée directement – le seul moyen d'y avoir accès est d'analyser, comme on l'a vu, les pratiques des agents – et on a vu que son maintien conduit Kuhn à soutenir des thèses très contestables sur le changement scientifique.

Dans la mesure où la notion de version, telle que je l'ai définie au chapitre 3, est elle aussi une entité hypothétique qui ne peut être étudiée que par ses manifestations (les pratiques des agents et les présentations explicites de la mécanique qu'ils proposent dans certains cas), on peut m'adresser la même objection : à quoi sert de maintenir

l'idée d'un ensemble de représentations mentales au principe des raisonnements et des pratiques des agents ? La différence majeure entre les versions comme ensembles de représentations mentales et les paradigmes comme cadres conceptuels partagés par tous les membres d'une communauté est que les premières concernent les individus : la notion de représentation mentale n'a, certes, pas de définition précise, mais c'est une hypothèse largement adoptée pour expliquer la vie cognitive des agents, qui ne souffre pas de la même imprécision que celle de paradigme.

Décrire l'activité théorique comme le développement et l'utilisation d'un ensemble de représentations mentales permet de rendre compte du fait, souligné ci-dessus, que les représentations théoriques acquièrent leur signification à la fois de leur utilisation dans la résolution de problèmes particuliers et de leur lien à d'autres représentations théoriques. À la lumière des analyses qui précèdent, supposer que chaque agent a sa propre version de la mécanique permet de comprendre qu'il puisse s'entendre avec ses pairs sur le sens des mots et sur l'usage qu'il convient de faire des équations de la mécanique sans pour autant supposer une identité complète de signification.

Au chapitre 3, j'ai examiné différentes versions de la mécanique, qui ont fait l'objet de présentations explicites au cours de l'histoire. J'ai alors souligné (section 2.2 du chapitre 3) que l'analyse des pratiques concrètes des agents est un autre moyen d'étudier leurs versions. L'analyse de la section 1.1 du présent chapitre m'a permis de montrer que l'étude de ces pratiques gagne à se concentrer sur l'utilisation de représentations particulières comme ce que Kuhn appelle « cas exemplaires ». La suite de ce chapitre vise ainsi à approfondir l'analyse de la compréhension *située* des agents au moyen d'unités d'analyse locales comme les cas exemplaires.

2 Les modèles

La notion de paradigme, dans son sens global de « matrice disciplinaire », s'est révélée reposer sur une hypothèse inutile dont le maintien a conduit Kuhn à adopter une conception holiste de la signification des théories qui contredit et annule ses considérations sur l'utilisation et la compréhension *située* des représentations théoriques par les agents. En son sens *local*, celui de *cas exemplaire*, la notion de paradigme permet en revanche d'attirer notre attention sur l'importance, dans l'apprentissage et la pratique de la science, de la résolution de problèmes particuliers – comme celui du pendule simple –, au moyen de laquelle les représentations théoriques, étant *appliquées* à ces problèmes, acquièrent une signification. Comme je l'ai mentionné au chapitre 2 (section 1.3.1), ces problèmes – et plus précisément les systèmes idéalisés comme le pendule simple qu'ils mettent en scène – sont très couramment appelés « modèles ».

Non seulement ces modèles, souvent présentés en guise d'exercice dans les manuels, jouent un rôle crucial dans l'*apprentissage*, mais encore, dans la mesure où les représentations théoriques – ce que Kuhn appelle « généralisations symboliques » – tirent leur signification de leur application à ces cas, ils sont constitutifs – et non pas seulement illustratifs – de la connaissance scientifique. Autrement dit, les lois de Newton, indépendamment de leurs applications à des cas typiques comme ceux de la chute libre d'un corps ou du pendule simple, n'ont pas, en elles-mêmes, de signification ; on ne les comprend qu'en tant qu'elles prennent la forme de représentations de systèmes particuliers – ou de types de systèmes.

C'est la raison pour laquelle plusieurs philosophes des sciences (par exemple Giere, 1988; Morgan et Morrison, 1999; Suárez, 1999; Cartwright, 1983, 1999; Frigg, 2002, 2006, 2009), au cours de ces dernières décennies, ont affirmé que l'analyse de l'activité et de la connaissance scientifiques gagnerait à se concentrer sur les *modèles* utilisés par les scientifiques, aussi bien dans l'apprentissage que dans la recherche.⁴⁵ Une des thèses partagées par ces différents auteurs est que la pratique scientifique ne consiste pas à *appliquer* des théories en suivant un schéma hypothético-déductif consistant à *tirer* les conséquences empiriques d'un ensemble d'axiomes, mais à *construire* des modèles, conçus comme des outils partiellement indépendants des théories, dont l'élaboration exige une grande part d'invention. Selon eux, ce sont les modèles, et non les théories, qui représentent effectivement les phénomènes et en permettent la prédiction et l'explication.

Cette attention aux modèles, fondée sur une critique des approches formelles des théories et sur le présupposé selon lequel la connaissance scientifique ne peut être correctement analysée que par une étude de la *pratique* de la science⁴⁶, se double, chez certains philosophes, de l'adoption d'une perspective *cognitive* : dès lors que l'on s'intéresse à l'action de représenter et à la compréhension que les agents acquièrent de la manipulation de représentations particulières, en appeler aux travaux

⁴⁵ Comme je l'ai signalé au chapitre 5 (note 51 page 212), dès les années 1960, plusieurs philosophes des sciences, comme Max Black (1962), Mary Hesse (1966) mais aussi, on l'a vu, Ernest Nagel (1961), soulignent l'importance des modèles dans l'activité scientifique. À la différence de ces philosophes (et en particulier de Nagel), ceux auxquels je m'intéresse dans la présente section sont en rupture complète avec les thèses de l'empirisme logique ; non seulement ils soulignent l'importance des modèles dans l'explication scientifique (comme Mary Hesse, par exemple), mais ils les présentent comme des unités d'analyse plus pertinentes que les théories, conçues comme des systèmes linguistiques.

⁴⁶ Ce tournant pragmatique en philosophie des sciences, qui consiste à étudier ce que les agents *font* et non les structures abstraites censées représenter les phénomènes, se manifeste dans les termes mêmes employés pour décrire la connaissance scientifique : ainsi par exemple, Ian Hacking (1983b) intitule son ouvrage « *Representing and intervening* », marquant, par l'usage du gérondif, le fait qu'il s'intéresse au « *représenter* » et non à « *la représentation* ».

des sciences cognitives sur les processus psychologiques qui gouvernent les représentations mentales, et sur la nature et la forme de ces dernières, semble pouvoir apporter un éclairage utile sur la nature de la connaissance scientifique elle-même.⁴⁷

Il doit être clair, à la lecture de ce qui précède, que j'emprunte ce tournant pragmatique et cognitif, que l'on peut décrire comme un double héritage kuhnien. Toutefois, les travaux des philosophes des sciences qui placent la notion de modèle au centre de leurs analyses souffrent souvent, selon moi, de plusieurs limitations, qui sont dues en particulier à une conception trop abstraite de la représentation. Dans cette section, je propose donc d'exposer les raisons pour lesquelles, tout en me situant dans la lignée de ces auteurs, je m'en démarque sur un point essentiel, en affirmant l'importance de la *forme particulière* des représentations utilisées par les agents, que ces philosophes négligent la plupart du temps.

2.1 Les différents sens du terme « modèle »

Avant de passer à l'analyse du rôle des modèles comme celui du pendule simple dans la pratique scientifique, il convient de dire un mot de la polysémie du terme « modèle », dans l'usage qu'en font les scientifiques aussi bien que les philosophes. On l'a, en effet, rencontré à plusieurs reprises au cours de ce travail, dans des sens souvent différents.

En un premier sens, un modèle est un *objet concret*, comme un schéma ou une maquette tridimensionnelle, qui sert à en représenter un autre. Le modèle de la mo-

⁴⁷Parmi les philosophes des sciences puisant dans les sciences cognitives des hypothèses permettant d'étudier les raisonnements à l'œuvre dans l'apprentissage mais aussi dans la découverte scientifiques, et, le cas échéant, d'offrir des outils d'analyse à l'histoire des sciences, on peut citer Nancy Nersessian (1984, 1992a,b, 2002a,b, 2007, 2008), Ronald Giere (1988, 2006), ainsi que Paul Thagard (1988, 1992). Voir aussi (Giere, 1992; Magnani *et al.*, 1999; Magnani et Nersessian, 2002). Ces travaux s'appuient pour la plupart sur l'hypothèse cognitive des *modèles mentaux*, selon laquelle les agents conduisent leurs raisonnements en manipulant des représentations qui ont la forme de modèles bidimensionnels de la réalité, et non pas (seulement) la forme séquentielle des représentations linguistiques (voir Gentner et Stevens, 1983; Johnson-Laird, 1983, 1996; Hirschfeld et Gelman, 1994). Ils empruntent également parfois des hypothèses aux travaux sur la cognition distribuée, selon lesquels le raisonnement ne consiste pas seulement en la manipulation de représentations mentales, mais s'appuie en grande partie sur des représentations externes, qui ne sont pas seulement la cause des représentations mentales, mais interviennent de façon essentielle dans les processus cognitifs (voir Hutchins, 1995; Clark, 1997). Ce domaine étant encore très nouveau, il est cependant difficile d'en tirer des résultats consensuels permettant d'appuyer l'analyse de la connaissance scientifique sur des hypothèses solides et bien définies; en conséquence, s'il semble bien que ce domaine soit fort prometteur pour une analyse de la connaissance scientifique dans une perspective cognitive, les références qui y sont faites dans les travaux de philosophie des sciences que j'ai mentionnés se limitent la plupart du temps à des remarques suggestives.

lécule d'ADN de Watson et Crick, le modèle réduit d'un avion ou d'un pont, offrent ainsi des représentations concrètes d'objets inaccessibles aux sens, ou impossibles à manipuler dans un laboratoire.

En un sens très différent, rencontré au chapitre 5, un modèle est une *structure mathématique* satisfaisant l'énoncé d'une théorie. Ici, il est plus difficile d'affirmer que le modèle est une *représentation*, du moins si l'on entend désigner par « représentation » une entité qui tient lieu d'une autre entité. La relation qu'entretiennent les axiomes de la théorie et le modèle est plutôt une relation de spécification ou de définition, dans un sens, et d'interprétation, dans l'autre : les axiomes spécifient (ou définissent) des structures, lesquelles sont des interprétations des axiomes.

Dans certaines de ses occurrences, le terme de modèle a un sens plus difficile à préciser : on l'a rencontré, chez Nagel, chez Kuhn et chez des physiciens comme Kelvin ou Boltzmann, pour désigner les *analogies* portées par l'utilisation d'un certain type de représentations, voire les *images mentales* que l'utilisation de ces représentations (externes) invite à se former. Chez ces mêmes auteurs, le terme de modèle sert aussi à désigner ces représentations (externes) elles-mêmes, qu'il s'agisse d'un type particulier d'équation, ou même d'un schéma concret représentant, par exemple, les molécules d'un gaz sous la forme de boules de billard.

Enfin, le terme de modèle sert parfois à désigner des *systèmes idéalisés* comme le pendule simple en mécanique, l'agent parfaitement rationnel en économie, ou encore la population isolée en génétique des populations. C'est à ce type de modèles que je m'intéresserai dans la section 2.2.

Je ne prétends naturellement pas que les différents sens que je viens de passer en revue (sans prétendre à l'exhaustivité) n'ont rien à voir entre eux ; cependant, chercher à forcer leur similitude d'entrée de jeu et, le cas échéant, vouloir abolir cette polysémie en restreignant le terme de modèle à certains de ses usages me semble une entreprise difficile et sans doute vaine. Difficile, parce qu'elle reviendrait à contraindre arbitrairement le langage en allant à l'encontre d'habitudes très ancrées. Vaine, parce que restreindre artificiellement notre usage du terme « modèle » ne serait ni suffisant ni même utile à résoudre les problèmes spécifiques posés par certaines choses qu'il sert à désigner.

Je propose donc simplement de prendre acte de cette polysémie, et de me concentrer à présent sur le dernier sens évoqué, celui où le terme « modèle » désigne des systèmes idéalisés comme le pendule simple. C'est sur le rôle de ce type de modèles dans l'apprentissage et dans la pratique scientifiques que les critiques des approches formelles, dont je voudrais à mon tour montrer les limites, ont le plus insisté. Ce faisant, on retrouvera, au cours de l'analyse, d'autres usages du terme de modèle, dont le lien et les différences avec celui-ci apparaîtront plus clairement que si l'on

avait cherché, d'entrée de jeu, à les préciser.

2.2 Les modèles comme « entités concrètes imaginaires » (Godfrey-Smith, 2006)

En insistant sur l'importance, dans la pratique scientifique, de l'utilisation de modèles comme ceux du pendule simple, les philosophes des sciences auxquels je m'intéresse dans cette section se présentent comme des critiques de la conception hypothético-déductive de l'activité scientifique véhiculée par l'empirisme logique, mais également, et surtout, de la conception sémantique des théories. Leur argument (en particulier chez Suárez, 1999 et Frigg, 2002, 2006, 2009) consiste en un refus de l'un des postulats les plus forts de la conception sémantique.

Ce postulat (voir chapitre 5, page 249) revient, précisément, à affirmer que les modèles utilisés par les scientifiques – comme celui du pendule simple, quand il sert à représenter un système particulier – peuvent être décrits comme des modèles logiques. Cela permet aux défenseurs de la conception sémantique de décrire l'application empirique des théories (voir chapitre 5, section 2.2) comme l'établissement d'une relation d'isomorphisme (ou d'une autre relation *entre* structures) entre les modèles théoriques – les structures qui satisfont les axiomes de la théorie – et les modèles physiques, au moyen desquels les scientifiques représentent les phénomènes. Ce faisant, ils prétendent parvenir à rendre compte, mieux que les empiristes logiques, de la pratique scientifique, dans la mesure où elle repose en grande partie sur l'utilisation de modèles.

Les critiques auxquelles je m'intéresse ici (Suárez, 1999; Frigg, 2002, 2006, 2009) affirment que cette analyse de l'application empirique des théories est aussi abstraite et éloignée de la pratique scientifique que celle des empiristes logiques.⁴⁸ Ils ne nient pas que l'on puisse décrire la relation entre les modèles théoriques – qui satisfont les axiomes de la théorie – et les modèles physiques comme le pendule simple comme une relation d'isomorphisme ou, plus généralement, comme une relation entre structures; en revanche, ils affirment que cela n'est pas suffisant pour rendre compte du processus par lequel les agents représentent les phénomènes au moyen de modèles physiques comme le pendule simple. Or, ils considèrent que le cœur du problème de

⁴⁸Ronald Giere (1988, 2006) est un cas à part, dans la mesure où il défend la conception sémantique des théories, mais propose une conception de l'utilisation des modèles par les agents qui s'apparente à celle défendue par les critiques de la conception sémantique, en cherchant, en outre, dans les résultats des sciences cognitives un moyen de rendre compte des processus cognitifs à l'œuvre dans l'utilisation de modèles comme ceux du pendule simple. Je ne présenterai pas sa position, mais elle correspond à peu de choses près, pour les questions qui m'intéressent ici, à celles de Frigg ou de Godfrey-Smith.

la représentation scientifique réside dans ce processus.

En effet, selon eux, on ne peut pas affirmer que les équations de la théorie, en tant qu'elles décrivent une structure, s'appliquent aux phénomènes. Décrire le comportement d'une horloge normande au moyen d'une équation du mouvement implique toujours un certain nombre d'idéalisations et d'approximations, qui consistent à ignorer, dans une certaine mesure, les frottements, la résistance de l'air, etc. Par conséquent, quel que soit son degré de précision, une équation du mouvement ne décrit pas, à strictement parler, le comportement de l'horloge normande, mais le comportement d'une « version idéalisée » (Frigg, 2009) de cette horloge.

Une grande part de l'activité scientifique, selon ces critiques des approches formelles, consiste en l'invention de ce type de versions idéalisées des phénomènes empiriques, que l'on appelle très couramment « modèles ». ⁴⁹ Cela consiste essentiellement en deux opérations (Godfrey-Smith, 2006; Frigg, 2009) : d'une part, *présenter* un système idéalisé hypothétique au moyen d'une description – comme un énoncé linguistique (« imaginons une masse ponctuelle accrochée dans le vide à un fil sans masse, etc. »), ou encore un schéma (comme celui de la figure 11) – qui permet d'écrire l'équation décrivant son comportement, et, d'autre part, *affirmer* que ce système idéalisé représente la portion du monde empirique que l'on cherche à étudier (par exemple, le mouvement du balancier d'une horloge normande).

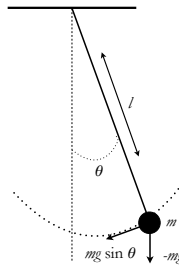


FIG. 11 – Schéma du pendule simple

Une analyse de l'activité scientifique doit donc, selon ces auteurs, être en mesure de rendre compte d'une part de la nature de ces systèmes idéalisés et d'autre part de la relation de représentation qu'ils entretiennent avec le monde réel. Contre la conception sémantique, ils affirment que l'on ne peut pas réduire ces systèmes idéalisés à des modèles au sens logique du terme, c'est-à-dire à des structures mathématiques.

⁴⁹ Je m'appuie, ici, sur l'exemple du pendule simple. Cependant, ces propos ne sont pas censés s'appliquer exclusivement à la physique : Godfrey-Smith (2006) montre ainsi que le travail de Levins en génétique des populations ainsi que les modèles de Maynard Smith et Szathmáry en théorie de l'évolution peuvent être considérés comme la description de populations hypothétiques idéalisées.

Le pendule simple ou la population isolée ne sont pas, selon eux, des structures, mais « des choses qui sont imaginaires ou hypothétiques, mais qui seraient *concrètes*⁵⁰ si elles étaient réelles. Une population imaginaire est une chose qui, si elle était réelle, serait une population en chair et en os, et non un objet mathématique. » (Godfrey-Smith, 2006, p. 735)

C'est, selon eux, en vertu de ce caractère « concret » que les agents sont en mesure de les utiliser comme des représentations de systèmes réels. C'est parce que le pendule simple est ce que Godfrey-Smith appelle une « entité concrète imaginaire » que l'on est en mesure de représenter une horloge normande *comme* un pendule simple. C'est parce que, *s'il était réel*, il ressemblerait à un pendule réel que les agents parviennent à décrire le balancier d'une horloge normande (un pendule réel) au moyen de l'équation du pendule simple. Cet acte ne consiste pas à appliquer une structure mathématique au balancier de l'horloge normande, mais à affirmer que l'on peut considérer le balancier d'une horloge normande *comme* un pendule simple, c'est-à-dire comme un objet constitué d'une masse ponctuelle suspendue à un fil sans masse. En termes kuhniens, c'est grâce à cette entité imaginaire qu'est le pendule simple, au moyen de laquelle ils comprennent la signification des équations de la mécanique, que les agents parviennent à voir les situations concrètes du monde empirique *comme* des situations newtoniennes, ce qui leur permet de les prédire et de les expliquer au moyen de ces équations.

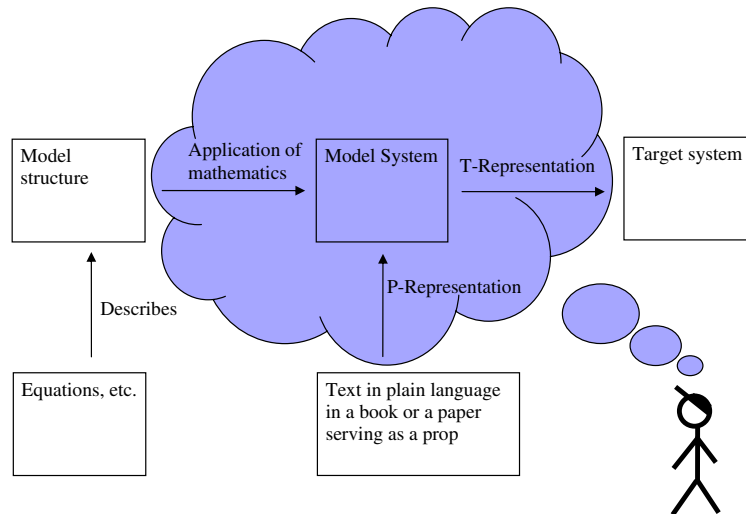


FIG. 12 – La représentation scientifique selon Roman Frigg (2009).

⁵⁰ Je souligne.

Ce faisant, ces auteurs proposent une conception de la représentation dans laquelle une entité imaginaire, qu'ils décrivent comme une fiction⁵¹, joue un rôle central. Pour en rendre compte, il convient de distinguer entre différentes relations de représentation et de description (voir le schéma qu'en propose Frigg, 2009, reproduit dans la figure 12). Prenons, à nouveau, l'exemple du pendule simple. D'une part, les équations du mouvement *décrivent* une structure mathématique, c'est-à-dire un modèle au sens logique (que Frigg appelle « *model structure* »). Entre cette structure et le système imaginaire qu'est le pendule simple (le « *model system* » dans le schéma de Frigg), est établie une relation mathématique comme, par exemple, un isomorphisme. Ce système imaginaire, lui-même, est *représenté* au moyen d'une description linguistique (un schéma peut aussi jouer ce rôle), tout comme Julien Sorel – qui n'existe pas – est décrit au moyen d'énoncés dans *Le Rouge et le Noir*. La relation entre l'énoncé linguistique et le modèle imaginaire est décrite comme une relation de représentation. Cette dernière (la « *P-representation* ») est cependant d'un autre ordre, selon Frigg, que celle (appelée « *T-representation* ») qui relie le modèle fictif – le pendule simple – au système réel – l'horloge normande.

Une analyse de l'activité scientifique doit, selon lui, être en mesure de rendre compte de ces différentes relations de représentation. Je voudrais à présent montrer pourquoi ce schéma, à supposer que l'on parvienne à résoudre les problèmes philosophiques posés par l'analyse de la fiction – ce que je ne chercherai pas à examiner – n'éclaire pas la manière dont les agents raisonnent au moyen des représentations qu'ils utilisent, parce qu'il reste tributaire d'une conception abstraite de la représentation.

2.3 L'importance de la forme particulière des représentations

Pour commencer, notons les deux avantages majeurs de l'analyse que je viens de présenter. En premier lieu, elle souligne le fait que les équations fournies par une théorie comme la mécanique classique n'acquiescent une signification que par leur application à des cas *concrets*, qu'ils soient imaginaires ou réels. Cela signifie, comme

⁵¹Peter Godfrey-Smith (2006, p. 735) affirme que, « bien que ces entités imaginaires soient déconcertantes, [...] elles peuvent être considérées comme similaires à quelque chose dont nous sommes tous familiers, les objets imaginaires de la littérature de fiction. » Roman Frigg (2009) développe cette conception des modèles comme fictions en s'appuyant sur les travaux de Kendall Walton (1990). Cette conception de la représentation scientifique sur le modèle de la fiction était déjà présente dans des travaux plus anciens en philosophie des sciences, comme ceux de Vaihinger (1911). Plus récemment, plusieurs philosophes des sciences ont trouvé dans les analyses de la fiction littéraire des outils féconds pour rendre compte de la représentation scientifique (voir Cartwright, 1983; Fine, 1993; Elgin, 1996; Barberousse et Ludwig, 2000). Voir aussi le volume consacré à ce sujet édité par Mauricio Suárez (2009). Ronald Giere (2009), en revanche, s'y oppose, et propose de concevoir les modèles comme des entités « abstraites ».

j'y ai insisté à maintes reprises, que c'est dans et par la pratique qui consiste à manipuler une équation pour résoudre des problèmes particuliers que celle-ci devient une authentique représentation. Autrement dit, on ne peut pas parler de représentation sans prendre en compte l'agent qui la comprend et l'utilise *pour représenter* un système concret – qu'il s'agisse du pendule simple ou d'une horloge normande. C'est la raison pour laquelle Godfrey-Smith (2006) et Frigg (2009) ont raison, selon moi, d'affirmer que le pendule simple est plus qu'une simple structure mathématique au sens où, s'il était réel, il aurait toutes les particularités d'un objet matériel. Cela ne revient à affirmer rien de plus que le fait que, quand j'imagine un objet tel que le pendule simple, je me le figure sous les traits d'un objet concret, qui n'est que sous-déterminé par la structure qu'il instancie ; ce faisant, je *me figure ce que dit la théorie* en tant qu'elle offre une représentation du monde empirique.

En second lieu, affirmer que les équations ne décrivent jamais – ou très rarement – parfaitement le comportement des systèmes réels permet d'attirer l'attention sur le fait que les idéalizations et les approximations nécessaires à la prédiction et à l'explication des phénomènes empiriques exigent de la part des scientifiques une grande capacité d'invention et d'imagination, qui, comme le souligne Kuhn, ne peut pas être simplement identifiée à la mise en œuvre de règles. Cela signifie donc qu'une grande part de l'activité scientifique n'est pas correctement décrite par la reconstruction formelle des théories, car ces dernières, ainsi reconstruites, ne nous disent pas comment effectuer ces opérations d'idéalisation et d'approximation.⁵² Cette compétence s'acquiert, comme le souligne Kuhn, par la résolution de problèmes de plus en plus compliqués, mettant en jeu des situations parfois inédites.

Une fois ces deux choses établies, il semble cependant que la relation que Frigg appelle « *T-representation* » (voir figure 12) est une relation abstraite, qui ne rend pas compte du processus par lequel, en pratique, les agents apprennent à utiliser les équations de la mécanique pour représenter des systèmes concrets (qu'ils soient idéalisés ou réels) et, ce faisant, acquièrent une compréhension de ces équations. Si l'on cherche à rendre compte de ce processus, on ne peut, comme je vais le montrer, le dissocier des inférences que les agents tirent au moyen de ces représentations. Cela apparaît clairement si l'on analyse en détail la résolution d'un problème de mécanique comme celui du pendule simple.

⁵²Comme je le signale au chapitre 5, page 258, cela ne me semble pas, pour autant, valoir comme une critique suffisante des approches formelles, qui ne prétendent pas (toujours) rendre compte de cet aspect de la pratique scientifique. Ma propre critique des approches formelles dans le chapitre 5 se situe en amont de cette critique, puisqu'elle consiste à montrer que l'analyse du contenu d'une théorie ne peut éviter de prendre en compte, dans une certaine mesure, la manière dont les agents l'utilisent.

J'ai montré, au chapitre 2 (section 1.3), que la résolution de ce problème implique plusieurs étapes. La première consiste à *identifier les forces* appliquées au système que l'on étudie afin d'*écrire l'équation de Newton sous la forme adéquate*. La deuxième consiste, une fois les conditions initiales du système connues, à résoudre cette équation. C'est en apprenant à reconnaître, dans un système concret, les forces qu'il faut prendre en compte pour en décrire le comportement, que j'acquiers la double compétence computationnelle et représentationnelle qui consiste, indissociablement, à savoir écrire correctement l'équation du mouvement et à savoir me représenter des situations de plus en plus complexes *comme* des situations newtoniennes. C'est ainsi, comme je l'ai souligné à plusieurs reprises, que j'acquiers une compréhension des équations newtoniennes.

Or, on l'a vu, la résolution du problème du pendule simple ne consiste pas dans les mêmes opérations quand j'utilise les équations de la mécanique analytique. Dans ce cas, écrire l'équation sous la forme adéquate ne consiste pas à identifier des forces, mais à choisir les coordonnées généralisées qui me permettent de trouver l'expression de l'énergie cinétique et de l'énergie potentielle du système afin d'en écrire le lagrangien. Manipuler les équations de Lagrange de façon à en faire les représentations d'une situation particulière implique de me représenter cette dernière *comme* une situation lagrangienne, c'est-à-dire une situation mettant en jeu des énergies, et non des forces.

Je ne conteste pas le fait que ce savoir-faire s'acquière d'abord en apprenant à manipuler les équations de la mécanique pour décrire des systèmes idéalisés comme le pendule simple, et que la maîtrise de ces cas simples guide, ensuite, l'utilisation des équations pour prédire et expliquer des situations du monde réel beaucoup plus compliquées. En revanche, je conteste le fait que la relation de représentation dont il s'agit de rendre compte pour comprendre l'activité scientifique soit celle que Frigg (2009) appelle la « *T-representation* », c'est-à-dire la « relation » entre le système idéalisé et les systèmes du monde empirique.

Qu'il s'agisse de l'application des équations de la mécanique à un cas imaginaire simple ou à un cas réel, les *représentations* que l'on manipule et au moyen desquelles on prédit et explique le comportement des systèmes (imaginaires ou réels) sont des équations dont la forme particulière (newtonienne ou lagrangienne) est cruciale dans les opérations que l'on doit effectuer. Représenter une situation du monde réel « au moyen du modèle du pendule simple » ne consiste pas dans les mêmes opérations selon que l'on utilise une représentation newtonienne ou une représentation lagrangienne. Autrement dit, on ne raisonne pas au moyen du modèle du pendule simple *in abstracto*, mais au moyen d'une représentation particulière de ce modèle. Cette représentation particulière, précisons-le, n'a pas toujours besoin d'être écrite noir sur

blanc sur un papier : je peux bien me contenter, dans les cas simples, de me représenter mentalement l'équation décrivant le mouvement d'un système, mais je me la représente toujours *sous une forme particulière*. Le pendule simple, *in abstracto*, n'est pas une représentation, si l'on entend par « représentation » un outil qui permet de tirer des inférences.

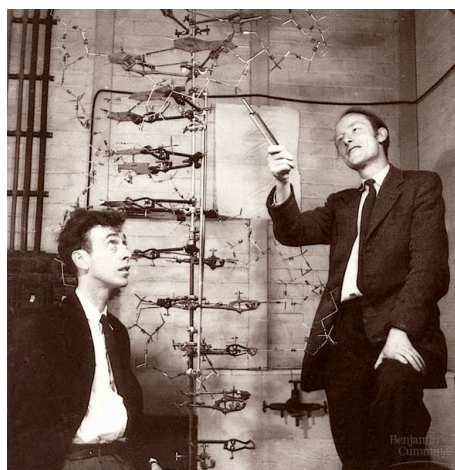


FIG. 13 – James Watson, Francis Crick, et le modèle de la double hélice de la molécule d'ADN

Cela devient clair quand on considère un autre type de représentation théorique, comme le modèle de la double hélice de l'ADN, proposé par Francis Crick et James Watson (voir figure 13). J'ai distingué, au début de cette section, les modèles de ce type de ceux qui, comme le pendule simple, sont des systèmes idéalisés. En effet, le modèle de la double hélice est un objet réel que Watson et Crick peuvent littéralement manipuler, alors que le pendule simple est imaginaire.

On peut, certes, raisonnablement affirmer que ce modèle matériel est la représentation d'une idéalisation. En effet, tout comme l'équation du pendule, il ne décrit à strictement parler aucune molécule d'ADN particulière, mais *exprime une hypothèse théorique* selon laquelle les molécules d'ADN peuvent être représentées sous la forme d'une double hélice. Si l'on se reporte au schéma de Frigg (2009), le modèle matériel concret sur lequel travaillent Watson et Crick joue le rôle des énoncés (« *text in plain language* ») qui décrivent le système imaginaire, mais n'est pas lui-même un système imaginaire. En effet, ce que l'on appelle *le* modèle de la double hélice n'est identifiable à aucun modèle matériel, mais est une hypothèse théorique, que l'on peut énoncer linguistiquement ou présenter sous la forme d'une maquette.

Pour autant, affirmer, en suivant le schéma de Frigg, que la maquette construite et manipulée par Watson et Crick n'est que le moyen d'accéder à une entité imaginaire,

qui elle-même sert à *représenter* les molécules réelles semble quelque peu coûteux et inutile. C'est bien la maquette réelle que Watson et Crick manipulent, et non pas une molécule d'ADN imaginaire. Qu'ils l'utilisent pour en tirer des informations sur les molécules d'ADN *en général* et non sur telle ou telle molécule particulière ne suffit pas à dire qu'ils utilisent une entité abstraite pour représenter des entités réelles.⁵³ Par là, je ne nie pas qu'il soit possible d'*imaginer* un objet tel que le pendule simple ; en revanche, j'affirme que l'on ne peut l'utiliser pour tirer des inférences – à propos des phénomènes en général ou de phénomènes particuliers – qu'en se le représentant (même mentalement) *sous une certaine forme*, et que cette forme joue un rôle central dans les inférences que l'on peut en tirer.

Ainsi, contrairement à ce que souligne Frigg (2009) dans son schéma de la représentation scientifique (figure 12), l'acte de représenter, par lequel les agents apprennent la signification des différentes équations de la mécanique et acquièrent les compétences nécessaires à leur utilisation, ne doit pas être étudié au moyen d'une enquête sur la nature de la *T-representation*, c'est-à-dire de la relation entre les idéalizations (*in abstracto*) et le monde réel, mais par une analyse détaillée des manipulations qu'impliquent la description des systèmes idéalisés et des systèmes réels au moyen de représentations particulières. Or, pour ces manipulations, la forme particulière de ces représentations est cruciale.

En conséquence, tout en reconnaissant l'importance des exercices qui consistent à appliquer ce que Kuhn appelle des « généralisations symboliques » à des cas idéalisés – que l'on peut appeler, suivant l'usage courant, des « modèles » –, et en admettant que la représentation des systèmes du monde réel repose sur des processus d'idéalisation – que l'on peut appeler, corrélativement, « modélisation » –, je suggère qu'une analyse des processus par lesquels les agents appliquent ces généralisations symboliques à des cas concrets (réels ou fictifs) doit prêter attention à la forme particulière de ces généralisations symboliques et au détail des manipulations et des inférences que cela implique. C'est ce que je me propose de faire dans la prochaine section.

⁵³C'est pourtant ce qu'affirme paradoxalement Ronald Giere (2006, pp. 105-106), après avoir souligné l'importance de la manipulation concrète de divers types de représentations – diagrammes, schémas, équations, etc. – par les agents : « l'expert utilise des représentations externes pour *re-construire* [italiques d'origine] des aspects du *modèle abstrait* [je souligne] pertinent pour l'étude d'un problème particulier. [...] Le modèle matériel de l'ADN de Watson et Crick, par exemple, sert à spécifier certains traits d'un modèle abstrait de l'ADN, comme la hauteur de l'hélice et les paires de bases autorisées. » Si, comme le prétend par ailleurs Giere, la relation de représentation repose sur des *ressemblances*, on voit mal en quoi le modèle abstrait ressemble à une molécule réelle. Si ressemblance il y a, elle est entre la maquette concrète et la molécule réelle. Je ne nie pas que la maquette concrète de la molécule d'ADN exprime une hypothèse théorique générale ; en revanche, c'est bien la maquette concrète qui sert de représentation et permet de tirer des inférences.

3 Unités d'analyse locales : figures et formules

On a vu, à la section 1, que l'étude de l'apprentissage et de la pratique scientifiques, en tant qu'elle nous renseigne sur la compréhension des représentations théoriques par les agents, gagnerait à se concentrer sur les unités d'analyse *locales* que sont les représentations que les agents utilisent et manipulent. En outre, j'ai montré qu'une analyse de leur signification devait prendre en compte la compréhension *située* qu'en ont les agents et renoncer à une sémantique vériconditionnelle, plutôt que d'en reconstruire le contenu logico-empirique *en principe*, cette dernière perspective ayant conduit Kuhn à tenir des propos qui contredisent ses considérations pourtant fécondes sur l'application des généralisations symboliques à des problèmes concrets.

L'analyse de la section 2 m'a permis de montrer que l'étude de l'interaction cognitive des agents avec ces unités d'analyse locales doit, de plus, prendre en compte leur forme particulière. En effet, si l'on prend au sérieux l'idée selon laquelle les représentations théoriques n'acquièrent une signification qu'en tant qu'elles sont appliquées, au moyen de manipulations concrètes, à des exemples particuliers, on ne peut ignorer la forme sous laquelle elles sont présentées aux agents, dans la mesure où ces manipulations en dépendent. Ainsi, utiliser le modèle du pendule simple ou de la double hélice de l'ADN pour représenter un système particulier consiste toujours à raisonner sur – et parfois littéralement manipuler – une équation, un diagramme, un graphe, ou encore un objet matériel particuliers.

Souligner et étudier l'importance de la forme particulière des représentations qui sont utilisés par les agents permet de rendre compte du caractère indissociable des deux compétences dont j'ai affirmé à plusieurs reprises qu'elles sont constitutives de la compréhension propre à l'activité théorique (voir chapitre 4, section 1.2) : celle qui permet aux agents de manipuler le « morceau de théorie » qu'ils utilisent pour résoudre un problème particulier de façon à lui donner la forme adéquate (que j'appelle *computationnelle*), et celle qui leur permet de voir une situation particulière comme, par exemple, une situation newtonienne (que j'appelle *représentationnelle*). La présente section vise à approfondir l'analyse de la relation entre ces deux dimensions de la compréhension, et en particulier à préciser la nature de celle que j'appelle « computationnelle ».

Pour cela, je propose dans un premier temps (section 3.1) de revenir sur la notion kuhnienne de « cas exemplaire » et sur l'analyse qu'il en fait dans les années 1970 (Kuhn, 1970a,b, 1969). On verra (section 3.1.1) que Kuhn souligne le caractère indissociable des aspects computationnels et représentationnels de l'activité théorique. Cependant, il minimise paradoxalement, dans certains de ses propos, l'importance de la dimension computationnelle ; je montrerai (section 3.1.2) que cela le conduit

à établir une frontière nette entre développement formel et invention théorique. Les limites de l'analyse de Kuhn me conduiront dans un deuxième temps (section 3.2) à présenter la notion de « gabarit », proposée par Paul Humphreys (2004). Cette notion est destinée, chez Humphreys, à souligner l'importance de la forme particulière – ce qu'il appelle la « syntaxe » – des équations utilisées dans la pratique scientifique ; ses propos me permettront d'approfondir l'analyse de la dimension computationnelle de l'activité théorique. Enfin, je conclurai cet examen (section 3.3) en montrant que, dès lors que l'on s'intéresse à la forme particulière des représentations et que l'on renonce à étudier leur signification sur le modèle de la référence linguistique, il est souhaitable de prêter attention à tous les types de représentations utilisés dans la pratique scientifique, et non pas seulement aux représentations linguistiques (comme les équations).

3.1 Les cas exemplaires : « Comment les scientifiques relient-ils les expressions symboliques à la nature ? » (Kuhn, 1969)

Si Kuhn, dans les années 1970, n'abandonne pas le sens global de la notion de paradigme, qu'il rebaptise « matrice disciplinaire », son insistance sur l'importance des cas exemplaires et l'analyse détaillée qu'il y consacre permettent d'y voir des unités d'analyse relativement indépendantes de leur insertion dans le cadre plus global de la matrice disciplinaire. C'est en tant qu'unités d'analyse locales autonomes que je vais m'y intéresser dans cette section.

Les étudier comme des unités locales autonomes revient à renoncer à l'idée selon laquelle elles tirent leur signification de la théorie de laquelle elles sont issues, et à les considérer comme des *formules* qui acquièrent une signification de leur application à des cas particuliers. Avant d'être utilisées pour résoudre un problème particulier, ces formules ne sont encore que des *généralisations symboliques*. Elles deviennent des *cas exemplaires* quand elles offrent la solution d'un problème particulier.

En effet, comme je l'ai signalé à la section 1 (page 290), ce que Kuhn appelle « cas exemplaire » n'est pas le système particulier – pendule simple ou horloge normande – auquel est appliquée la généralisation symbolique ; c'est la généralisation symbolique elle-même, une fois mise sous la forme particulière qui permet à l'agent de résoudre le problème. C'est, comme le dit Kuhn, la « solution concrète d'un problème » [*concrete problem solution*]. Autrement dit, c'est la généralisation symbolique *en tant qu'elle est interprétée*. Il convient à présent d'aborder de front l'analyse de ce processus d'interprétation, par lequel une généralisation symbolique devient un cas exemplaire.

3.1.1 Manipulation et interprétation des généralisations symboliques

On a vu, à la section 1 (page 289), que Kuhn refuse l'idée selon laquelle l'application des généralisations symboliques à la résolution d'un problème particulier consisterait en la mise en œuvre d'un ensemble de *règles*. Une des preuves, selon lui, de l'inadéquation d'une telle description est qu'elle ne rend pas compte de la capacité des scientifiques à écrire les généralisations symboliques sous des formes parfois inédites, adaptées à la résolution d'un nouveau type de problème.

Cette capacité, comme tout savoir-faire, s'acquiert par l'exercice. C'est en ayant successivement appris à appliquer l'expression symbolique du principe fondamental de la dynamique, « $f = ma$ », à la chute libre d'un corps, au mouvement d'un pendule simple et à une paire d'oscillateurs harmoniques couplés, problèmes pour lesquels elle prend des formes très différentes (« $mg = md^2s/dt^2$ » pour la chute libre, « $mg\sin\theta = -md^2s/dt^2$ » pour le pendule simple, et « $m_1d^2s_1/dt^2 + k_1s_1 = k_2(d+s_2-s_1)$ » pour la paire d'oscillateurs harmoniques couplés) qui la rendent méconnaissable aux yeux du novice, que l'étudiant acquiert le savoir-faire qui en fait le membre d'une communauté scientifique (Kuhn, 1969, p. 465).

Ce savoir-faire, s'il consiste bien à manipuler une expression symbolique, ne se réduit pas, nous dit Kuhn, à une procédure « purement syntaxique » :

Avant de pouvoir commencer les manipulations logiques et mathématiques qui aboutissent à la prédiction de mesures précises, le scientifique doit écrire la forme particulière de $f = ma$ qui s'applique, par exemple, à la vibration d'une corde ou la forme particulière de l'équation de Schrödinger qui s'applique par exemple à l'atome d'hélium dans un champ magnétique. Quelle que soit la procédure qu'il met en œuvre en faisant cela, elle ne peut pas être purement syntaxique. (Kuhn, 1969, p. 466)

Ce que Kuhn suggère ici – quoique ce ne soient pas les termes qu'il emploie –, c'est qu'il convient de distinguer entre *deux types de compétences computationnelles*, qui correspondent à deux types de manipulation de la forme des équations. La première, qu'il appelle « purement syntaxique », consiste en la mise en œuvre des règles de calcul logico-mathématiques qui permettent, par exemple, de trouver les solutions d'une équation différentielle. Ces calculs ne requièrent pas, en tant que tels, l'interprétation de ces expressions. Cette capacité à manipuler des expressions non interprétées est la seule partie de la compétence scientifique qui puisse, selon Kuhn, être décrite comme l'application de règles.

La seconde, celle qui permet d'écrire l'équation de Schrödinger sous la forme particulière qui s'applique à l'atome d'hélium dans un champ magnétique, consiste aussi en la manipulation d'une expression symbolique, mais elle n'est possible qu'en tant que cette expression est *interprétée*. Ce qui guide ce deuxième type de manipulation,

c'est donc la *compréhension de la signification de cette expression*.⁵⁴

Ainsi, contrairement à ce que suggère l'image empiriste logique des théories qui les décrit comme des ensembles d'énoncés syntaxiques joints à des règles d'interprétation, l'interprétation ne vient pas *après coup* : quand l'expression « $f = ma$ » est transformée, par exemple, en « $mg\sin\theta = -md^2s/dt^2$ »,

« $f = ma$ » n'est plus une expression non interprétée. Le processus qui consiste à adapter les cas exemplaires aux expressions est *initialement*⁵⁵ une manière d'apprendre à interpréter ces expressions. Quand on peut le faire, on a appris l'interprétation. (Kuhn, réponse à Achinstein, in Suppe, 1974/1977b, p. 517)⁵⁶

Être capable d'écrire une équation de la forme « $mg = md^2s/dt^2$ » pour un problème de chute libre, « $mg\sin\theta = -md^2s/dt^2$ » pour un pendule simple, ou encore « $m_1d^2s_1/dt^2 + k_1s_1 = k_2(d + s_2 - s_1)$ » pour une paire d'oscillateurs couplés, c'est déjà *comprendre* « $f = ma$ » d'une certaine manière. C'est cette compréhension de l'équation que l'étudiant acquiert au moyen de cas exemplaires.

Les étudiants en physique affirment régulièrement avoir lu l'intégralité d'un chapitre de leur manuel et l'avoir *parfaitement compris*⁵⁷, mais avoir malgré cela des difficultés à résoudre les problèmes à la fin du chapitre. Presque toujours, ces difficultés résident dans l'écriture des équations appropriées et dans la mise en relation des mots et des exemples donnés dans le manuel aux problèmes particuliers qu'on leur demande de résoudre. Ordinairement, ces difficultés se dissolvent toutes de la même manière. L'étudiant découvre une manière de voir ce problème comme un problème qu'il a déjà rencontré. Une fois que cette similitude ou analogie est perçue, demeurent seulement les difficultés de manipulation. (Kuhn, 1969, p. 470)

Les deux types de « difficultés » mentionnés ici correspondent aux deux types de manipulations que j'ai distingués : celles qui consiste à écrire les équations sous la forme adéquate disparaissent quand l'étudiant comprend ces équations comme les représentations d'un certain type de systèmes ; les secondes, les seules que Kuhn appelle ici des « difficultés de manipulation », concernent, elles, la maîtrise de règles

⁵⁴On avait déjà rencontré, chapitre 4 (page 168), cette opposition entre une compétence que Kuhn appellerait « purement syntaxique » et que Feynman attribue aux mathématiciens et une compétence, computationnelle aussi (au sens où elle consiste en la manipulation de la forme des équations), qui est guidée par l'interprétation ou l'« intuition » physique.

⁵⁵Je souligne.

⁵⁶Cette phrase est extraite de la réponse à une critique d'Achinstein lors de la conférence sur la structure des théories scientifiques qui s'est tenue dans l'Urbana en 1969 (Suppe, 1974/1977b), où Kuhn a présenté ses « Second thoughts on paradigms ». Achinstein (Suppe, 1974/1977b, p. 516) reprochait à Kuhn de concéder trop à l'empirisme logique en supposant que, avant sa transformation, « $f = ma$ » est non interprétée.

⁵⁷Je souligne.

logico-mathématiques. C'est la maîtrise des premières, qui sont indissociables d'un processus d'interprétation, et non des secondes, qui détermine le plus précisément, selon Kuhn, l'appartenance à une communauté scientifique.

En effet, alors que « beaucoup de communautés scientifiques partagent, par exemple, l'équation de Schrödinger », elles ne l'utilisent pas toujours pour résoudre le même type de problèmes. Par conséquent, les spécialistes de la physique des solides et ceux de la théorie des champs, par exemple, ne « partagent de manière non équivoque » que « l'équation de Schrödinger *non interprétée*⁵⁸, et non pas une équation interprétée » (Kuhn, 1969, p. 471, note 17).

Ainsi, Kuhn rejette fermement l'idée selon laquelle les scientifiques disposeraient d'une batterie de généralisations symboliques abstraites comme, par exemple, « $f = ma$ » ou l'équation de Schrödinger, auxquelles ils donneraient *ensuite* une interprétation empirique en en reliant les différents éléments à des éléments du monde.⁵⁹ Pour le scientifique qui l'utilise, l'expression de la deuxième loi de Newton, *même à un très haut niveau de généralité*, c'est-à-dire avant même de servir à la description, par exemple, du mouvement d'un pendule, est interprétée.⁶⁰ C'est la raison pour laquelle Kuhn affirme que « les cas exemplaires (et aussi les modèles) sont des déterminants bien plus efficaces de la sous-structure des communautés que ne le sont les généralisations symboliques. » En effet, dès lors qu'elles sont utilisées, les généralisations

⁵⁸ Je souligne.

⁵⁹ Signalons que Kuhn, au moment de l'édition, par Frederick Suppe (1974/1977b), des actes de la conférence tenue en 1969, ajoute une note soulignant une confusion importante à propos de la notion de règles de correspondance dans sa propre contribution à la conférence. En effet, dans sa critique de la conception positiviste logique de l'interprétation empirique des théories, il emploie à tort l'expression « règles de correspondance » ou « règles d'interprétation » pour désigner le processus par lequel les termes théoriques sont « attachés à la nature ». Or, comme il le souligne dans sa note, « les règles de correspondance relient uniquement des mots à d'autres mots, et pas à la nature. » Il ajoute qu'il prétend aussi bien critiquer l'idée selon laquelle les termes théoriques sont reliés aux termes observationnels au moyen de règles de correspondance, que celle, dont il affirme que les empiristes logiques en font une hypothèse implicite, selon laquelle « tout individu qui sait utiliser correctement un terme de base a accès, consciemment ou inconsciemment, à un ensemble de critères qui définissent ce terme ou fournissent les conditions nécessaires et suffisantes pour son application. » « Pour cette sorte d'attachement-par-critère j'ai aussi utilisé l'expression "règle de correspondance", ce qui est une violation de son usage normal. Cette extension est cependant excusable dans la mesure où je crois qu'en appeler explicitement aux règles de correspondance et implicitement à des critères d'application revient à introduire la même procédure et attire de la même manière notre attention dans une mauvaise direction. » (Kuhn, 1969, p. 467, note 11)

⁶⁰ C'est ainsi que je propose de comprendre l'affirmation de Kuhn selon laquelle « le contenu empirique doit entrer dans les théories formalisées par le haut aussi bien que par le bas » et que les « formalismes en science sont reliés à la nature par le haut, sans l'intervention d'une déduction qui élimine les termes théoriques. » (Kuhn, 1969, p. 466).

symboliques sont des cas exemplaires.

Quand je dis que [les membres d'une communauté] partagent un engagement pour, par exemple, la généralisation symbolique $f = ma$, je veux seulement dire qu'ils ne font aucune objection à un homme qui inscrit les quatre symboles, f , $=$, m , et a à la suite sur une ligne, qui manipule les expressions résultantes au moyen de la logique et des mathématiques, et qui exhibe un résultat lui-même symbolique. Pour nous, à ce stade de la discussion, *quoique pas pour les scientifiques qui les utilisent*⁶¹, ces symboles et les expressions qui résultent de leur composition sont non interprétés, encore vides de signification empirique ou d'application. Un engagement partagé pour un ensemble de généralisations justifie des manipulations logiques et mathématiques, et implique un engagement pour le résultat. Cela n'implique pas obligatoirement, cependant, un accord sur la manière dont ces symboles, individuellement ou collectivement, sont reliés aux résultats de l'expérience et de l'observation. Dans cette mesure, les généralisations symboliques partagées fonctionnent comme des expressions d'un système purement mathématique. (Kuhn, 1969, p. 464)

L'« engagement » partagé par tous les physiciens qui utilisent l'équation de Schrödinger se réduit à l'acceptation commune de la relation logique des symboles qui y figurent et à celle des règles de calcul de la logique et des mathématiques, qui permettent de tirer les conséquences de cette équation, quels que soient les phénomènes qu'elle sert à représenter. La mise en œuvre de ces règles de calcul ne requiert aucune connaissance de la nature des référents de ces symboles. Cependant, si elle est encore vide de signification pour l'observateur extérieur qui constate que tous les individus appelés « physiciens » utilisent cette suite de symboles, elle est toujours déjà interprétée pour les scientifiques qui l'utilisent, sans quoi ils seraient incapables de l'écrire sous la forme idoine (tout comme le débutant qui n'a pas encore acquis cette compétence).

En conséquence, selon les cas auxquels ils l'appliquent, les membres de communautés scientifiques différentes ne donnent pas la même signification à cette « même » généralisation symbolique, ce qui permet à Kuhn, comme on l'a vu dans la première section de ce chapitre (page 291), d'affirmer que, « bien que la mécanique quantique (ou la dynamique newtonienne, ou la théorie électromagnétique) soit un paradigme pour de nombreux groupes scientifiques, *ce n'est pas le même paradigme pour tous*⁶² » (Kuhn, 1962/1970, pp. 49-50).

Distinguer ainsi entre deux compétences que l'on peut appeler « computationnelles », au sens où elles se manifestent toutes deux par la manipulation d'une expression symbolique, permet de souligner le fait que celle qui caractérise la compré-

⁶¹ Je souligne.

⁶² Je souligne.

hension est indissociable d'un processus d'interprétation, et donc de la compétence que j'appelle « représentationnelle ». Dire qu'elles sont indissociables, c'est refuser le présupposé, qui sous-tend l'analyse des empiristes logiques, selon lequel la forme des énoncés et leur contenu peuvent être clairement distingués.

Ajoutons pour finir que cette conviction kuhnienne, que je reprends à mon compte, selon laquelle il est impossible de distinguer le processus de manipulation d'une expression symbolique de celui de son interprétation, est à mon sens indépendante du holisme sémantique qu'il adopte par ailleurs et qui le conduit à soutenir la thèse de l'incommensurabilité des paradigmes. On peut en effet affirmer que l'étude de l'activité théorique interdit de distinguer nettement, au niveau local de l'utilisation des représentations issues des théories, entre des formules non interprétées et le contenu de l'observation, sans pour autant adopter l'idée d'un schème conceptuel qui détermine, globalement, une certaine vision du monde.

Pourtant, tout en distinguant entre ces deux compétences computationnelles, et en remarquant que les processus d'interprétation et de manipulation des généralisations symboliques sont indissociables, Kuhn minimise, dans certains de ses propos, l'importance de la dimension computationnelle du processus d'interprétation. Cela se manifeste, comme on va le voir, dans la distinction nette qu'il trace entre le développement formel des théories et l'invention théorique.

3.1.2 Invention de nouveaux « formalismes » et invention théorique

Comme on vient de le voir, écrire une équation sous la forme adaptée à un problème particulier ne consiste pas seulement à appliquer des règles, mais exige une capacité d'*invention* de la part de l'agent, laquelle repose, pour ainsi dire, sur l'intuition physique qu'il a acquise en résolvant des problèmes de plus en plus compliqués. En effet, comme le note Kuhn, « chacune de ces classes [de problèmes] exige un nouveau formalisme. » Pourtant, Kuhn, comme je vais le montrer dans cette section, en ne procédant pas à une analyse précise de ce qu'il faut entendre ici par « formalisme », minimise l'importance de la forme même des représentations manipulées par les scientifiques, et fait, à nouveau, une concession majeure à l'empirisme logique.

Avant de montrer en quoi cette concession consiste, soulignons certaines conséquences de la description de l'activité théorique présentée dans la section précédente. Dans la mesure où l'activité des scientifiques ne peut pas être décrite comme l'application de règles, l'entreprise des empiristes logiques consistant à expliciter les règles de déduction et d'interprétation qui caractérisent l'application empirique des théories, ne rend pas compte des processus effectifs qui sont mis en œuvre par les scientifiques. Parce que ce sont, comme on l'a vu, ces processus effectifs qui déterminent la si-

gnification des représentations ainsi manipulées, Kuhn affirme que la description des théories comme des ensembles de formules non interprétées auxquelles sont ajoutées des règles de correspondance conduit à « des reconstructions de théories légèrement différentes » (Kuhn, 1969, p. 468) de celles qui sont effectivement utilisées par les scientifiques.

Ce que Kuhn semble vouloir dire ici, c'est que la reconstruction formelle d'une théorie, *si on la considère comme une formulation de cette théorie*, exprime en fait une théorie différente : si l'on suppose que les règles qui y sont explicitées décrivent des processus d'interprétation, alors la théorie reconstruite n'est pas la théorie telle qu'elle est effectivement utilisée et comprise par les scientifiques, mais une théorie qui *pourrait être utilisée*. En conséquence, la reconstruction formelle, plutôt qu'une analyse de la connaissance scientifique, s'apparente plutôt aux travaux des scientifiques eux-mêmes, quand ils cherchent à développer formellement leur théorie pour la rendre plus puissante ; c'est du moins ce que Kuhn suggère quand il affirme que le « philosophe peut bien, en se comportant comme un scientifique, améliorer la théorie d'un groupe, mais il ne l'aura pas, en tant que philosophe, analysée. » (Kuhn, 1969, p. 468)

Cela signifie, comme le suggère Suppes (1968) que Kuhn cite en note comme un « exemple typique de l'hypothèse selon laquelle un perfectionnement du formalisme est ipso facto une analyse de la connaissance déployée par la communauté qui utilise ce formalisme à améliorer » (Kuhn, 1969, p. 468, note 13), qu'explicitier les hypothèses fondamentales et la structure d'une théorie peut contribuer à l'améliorer et à la rendre plus puissante, mais que cela ne contribue pas à éclairer l'activité théorique effective consistant en l'utilisation de la théorie sous sa formulation – peut-être imparfaite – actuelle.⁶³

Or Kuhn compare explicitement la tâche de reconstruction formelle à celle qu'a accomplie Hamilton en mécanique classique :

⁶³ Ainsi Kuhn affirme-t-il que « les formalistes ont souvent confondu la tâche consistant à *améliorer* la clarté et la structure des éléments formels d'une théorie scientifique avec la tâche très différente consistant à *analyser* la connaissance scientifique. » (Kuhn, 1969, p. 468, note 13) Dans sa réponse à la critique de Suppes, Kuhn ajoute : « Vous essayez [...] d'améliorer les théories – les scientifiques le font et il n'y a pas de raison pour laquelle un philosophe ne devrait pas le faire – et je ne veux pas tracer de frontière qui vous exclue. Pour votre problème, qui plus est, je pense que vous avez raison : plus une théorie scientifique est formalisée, plus elle a de chances d'être puissante. Mais la théorie formalisée n'est pas la même théorie, et vous n'avez pas, en l'améliorant, simplement trouvé ce qui était implicite dans ce que les gens faisaient avant. L'épistémologue doit donc être un peu méfiant à l'idée de l'accepter comme une version explicite de la théorie qu'il avait auparavant. Si les théories peuvent fonctionner sans autant de formalisme, il doit comprendre comment et pourquoi, et non pas produire une version plus formalisée. » (Kuhn, 1969, p. 513)

Hamilton a produit une meilleure formulation de la mécanique newtonienne que celle de Newton, et le philosophe peut espérer qu'une formalisation accrue puisse déboucher sur des progrès. Mais il ne doit pas prendre pour acquis qu'il en ressort la même théorie que celle de laquelle il est parti, ni que les éléments formels de l'une ou l'autre version de la théorie sont coextensifs avec la théorie elle-même. » (Kuhn, 1969, p. 468, note 13)

Ces propos semblent clairs : les développements formels de Hamilton conduisent à une théorie dont l'application – et donc, serait-on tenté de conclure – la signification, n'est pas la même que celle formulée en termes newtoniens. En effet, si la signification d'une théorie pour un groupe donné dépend de la manière dont il l'utilise, ce qui implique le choix d'un certain type d'expression symboliques et leur application *via* un certain type de procédures interprétatives, alors on doit pouvoir conclure que la mécanique sous ses différentes formulations n'a pas la même signification. En ce sens, l'invention de ce que Kuhn appelle des « formalismes » nouveaux, parce qu'ils permettent des applications nouvelles et s'accompagnent, comme on l'a vu, d'une nouvelle manière de relier les équations de la mécanique aux phénomènes empiriques, devrait compter comme une authentique invention théorique.

Pourtant, et c'est ici que se situent, selon moi, les limites de l'analyse de Kuhn, il insiste paradoxalement, à plusieurs reprises, sur la différence entre le développement formel d'une théorie *au sein d'un paradigme* et l'invention d'une « nouvelle » théorie. C'est en effet ce même exemple – celui du développement « formel » de la mécanique classique par Hamilton – que Kuhn convoque pour souligner que l'écriture d'un « formalisme spécial, d'une nouvelle version de la formalisation, ne peut pas consister dans la même chose que l'invention d'une nouvelle théorie. Entre autres choses, [l'écriture d'un formalisme spécial] peut être enseignée alors que l'invention d'une théorie ne peut pas l'être. » (Kuhn, 1969, p. 466)

En dressant ainsi une barrière entre la compétence, acquise par les étudiants au moyen de cas exemplaires, qui leur permet d'écrire une généralisation symbolique sous la forme adaptée à différents problèmes, et celle qui permet d'« inventer une nouvelle théorie », Kuhn minimise ce qu'il a pourtant contribué à souligner : le fait que la manipulation de généralisations symboliques ne consiste pas en l'application de règles, mais requiert une part d'invention. Alors que ses propos sur l'utilisation des généralisations symboliques permettent de rendre compte de multiples différences au sein d'une communauté large de chercheurs utilisant la « même » généralisation symbolique, dresser une telle frontière contribue, une nouvelle fois, à « rigidifier » la notion de paradigme et à admettre que tout changement théorique est un changement logico-empirique.⁶⁴

⁶⁴ Comme le remarque Anouk Barberousse (2008, p. 97), « la thèse de Kuhn, qui émerge de l'op-

Finalement, contrairement à ce qu'il affirme ailleurs⁶⁵, Kuhn maintient implicitement l'idée selon laquelle, « au sommet » des théories, les généralisations symboliques – qui peuvent recevoir plusieurs interprétations différentes selon le domaine de phénomènes auxquelles elles sont appliquées – sont des unités fixes et interchangeables. Chacun des groupes utilisant la formule « $f = ma$ » est enclin à la manipuler et à l'interpréter d'une certaine manière, mais la forme mathématique elle-même de cette équation ne semble pas jouer de rôle particulier dans ce processus computationnel.

À vouloir, comme on l'a vu, distinguer le processus de manipulation des expressions symboliques issues de la théorie de celui d'une manipulation purement syntaxique – en distinguant deux types différents de « manipulations » –, Kuhn est paradoxalement conduit à ignorer un aspect pourtant très important de ce processus : le fait que la forme particulière de la généralisation symbolique joue un rôle crucial – c'est-à-dire, le fait que, comme Paul Humphreys (2004) le souligne, la « syntaxe » compte. Autrement dit, Kuhn ignore la dimension théorique que peut avoir un changement formel. Cela se manifeste, d'une part, dans les propos qu'il tient sur les développements « purement formels » opérés par Hamilton⁶⁶ et, d'autre part, dans le fait qu'il ne prend jamais en considération, dans ses exemples, des cas de généralisations symboliques « équivalentes » présentées sous des formes différentes, comme par exemple les équations de la mécanique. En effet, les exemples de Kuhn sont toujours ceux d'une généralisation symbolique – la loi de Newton, l'équation de Schrödinger – et des différentes formes qu'elle prend dans la pratique.

Probablement en raison même de son rejet de l'empirisme logique, qui véhicule

position explicite entre des “formalismes spéciaux” et la “théorie”, est que l'invention d'une nouvelle théorie est nécessairement causée par une découverte empirique, contrairement à l'écriture de nouveaux formalismes. »

⁶⁵Voir note 60.

⁶⁶C'est un exemple qu'il prend à plusieurs reprises, notamment dans un article portant sur la méthode de reconstruction formelle des théories proposée par Sneed (1971). Cette méthode, développée également par Stegmüller, et exposée par Balzer *et al.* (1987), est une méthode structuraliste proche de la méthode ensembliste de Suppes, qui cherche cependant à rendre compte de la dynamique de la science telle qu'elle est décrite par Kuhn ainsi que par Lakatos. Kuhn (1976), tout en affirmant sa sympathie pour cette méthode, regrette cependant que, telle quelle, elle ne permette pas encore de répondre à la question suivante : « Qu'est-ce qui sous-tend notre conviction selon laquelle la mécanique relativiste diffère conceptuellement de la mécanique newtonienne en un sens où, par exemple, les mécaniques lagrangienne et hamiltonienne n'en diffèrent pas ? » (Kuhn, 1976, p. 189 ; 2000, p. 187) Dans une note de ce même article, il affirme que, quoique « les relations entre les trois formulations de la mécanique classique du point soient plus complexes [que celles de la mécanique du solide et de la mécanique du point], elles ont coexisté sans être considérées incompatibles. Il n'y a pas de raison apparente de supposer que l'introduction d'une mécanique autre que newtonienne ait constitué une révolution. » (Kuhn, 1976, p. 198, note 10 ; 2000, p. 188, note 18)

une conception *syntaxique* des théories, Kuhn, se méfiant, pour ainsi dire, de la « syntaxe », rejette toute considération accordant de l'importance à la forme même des représentations pour les inférences des agents. Ce faisant, il rend assez mystérieuse cette capacité à tirer des prédictions à partir de généralisations symboliques, qu'il a l'air de considérer comme données une fois pour toutes au sommet de la théorie.⁶⁷ En ignorant le détail des inférences effectivement faites par les agents, il ne permet pas de les distinguer clairement de l'application de règles décrivant tout ce qui peut, en principe, être déduit d'une formule.

En cherchant à ne rien concéder à l'empirisme logique, Kuhn lui fait, une nouvelle fois, une concession majeure : il admet implicitement que la forme particulière des représentations est sans importance. C'est ainsi que je propose de comprendre son refus de prendre les différences entre équations newtoniennes, lagrangiennes et hamiltoniennes au sérieux, alors que plusieurs éléments de son analyse semblent y inviter. Cette concession repose, selon moi, sur une confusion entre deux manières de comprendre la notion de « syntaxe », que la prochaine section vise à éclairer.

3.2 Les gabarits : « La syntaxe compte. » (Humphreys, 2004)

Dans le chapitre intitulé « Computational Science » de *Extending Ourselves*, Paul Humphreys (2004) définit une nouvelle unité d'analyse de l'activité théorique, les *gabarits*.⁶⁸ Elle est précisément destinée à rendre compte de l'importance de la forme particulière – que Humphreys appelle « syntaxe » – des équations et, plus généralement, des schèmes de calculs mis en œuvre dans la prédiction et l'explication des phénomènes.⁶⁹ Comme on va le voir, l'usage fait par Humphreys du terme « syntaxe »

⁶⁷C'est d'ailleurs une critique semblable que lui fait Paul Humphreys (2004, p. 101), à propos du caractère « routinier » de l'utilisation des généralisations symboliques par les membres d'une communauté : « il est étonnant que Kuhn, qui savait que l'on n'apprend rien concernant la science en regardant uniquement ses produits finis, ait été aussi attaché aux méthodes des manuels dans ce cas. L'apprentissage et l'utilisation routinière de modèles tous faits n'est certainement pas caractéristique de l'usage tranchant des modèles de recherche, même au sein de la science normale. »

⁶⁸Le terme anglais est « *templates* ». Utilisé dans plusieurs domaines différents, ce terme peut servir à désigner les *formes*, au sens de « moules », utilisées dans la fabrication d'objets, ou encore les *patrons* utilisés dans la confection de vêtements. En informatique, ce terme désigne un modèle syntaxique de conception de logiciel ou de présentation des données ; il est aussi utilisé en linguistique, pour désigner une unité syntaxique (voir Lowenstamm, 2003). Dans ces deux usages, la traduction française de « *template* » est « gabarit ». Dans la mesure où le terme est choisi par Humphreys pour insister sur la dimension syntaxique des équations utilisées dans la pratique scientifique, c'est l'analogie avec l'informatique et la linguistique qu'il m'a semblé le plus pertinent de souligner.

⁶⁹Signalons que les propos de Humphreys s'insèrent dans une entreprise visant à rendre justice à l'importance grandissante de l'utilisation des ordinateurs dans la pratique scientifique, ces derniers augmentant nos capacités de calcul au point, selon lui, d'appeler une redéfinition de la tripartition

n'implique pas un retour à une conception syntaxique abstraite des théories ; bien au contraire, il vise à montrer que des équations dont on peut en principe déduire les mêmes conséquences en vertu d'un ensemble de règles logiques ne permettent pas toujours les mêmes opérations, en pratique. Ces opérations dépendent en effet de la forme sous laquelle l'équation est présentée et des limitations des processeurs – ou des agents cognitifs – qui les implémentent.

3.2.1 Les « morceaux concrets de syntaxe »

Humphreys propose d'appeler « gabarits » les différents schèmes de calcul ou algorithmes utilisés dans la pratique scientifique. Ceux-ci incluent des formes d'équations, comme par exemple les équations lagrangiennes, des modèles statistiques comme les distributions de Poisson, mais également des outils théoriques comme les automates cellulaires nés des travaux de von Neumann sur les systèmes auto-réplicatifs (von Neumann et Burks, 1966).

Prenons l'exemple d'une équation lagrangienne. Indépendamment de son contenu représentationnel et du domaine de phénomènes que la théorie dont elle est issue est censée expliquer⁷⁰, elle se caractérise par une certaine syntaxe et conséquemment par un certain type de manipulation et de calcul que les scientifiques doivent effectuer pour en tirer des prédictions. C'est cette dimension syntaxique de l'équation, que l'on peut utiliser pour décrire différents types de phénomènes, que Humphreys appelle « gabarit ».⁷¹

Ainsi, une équation newtonienne et une équation lagrangienne décrivant le même système, ou encore une équation en coordonnées rectilignes et la « même » équation en coordonnées polaires, ont des gabarits différents. Bien que, représentant le même

classique entre théorie, observation et expérimentation (Humphreys, 2004, p. 51). Ma perspective est donc assez différente de la sienne, puisque je m'intéresse en premier lieu aux inférences des agents cognitifs (et non aux opérations effectuées par les ordinateurs) ; toutefois, les réflexions de Humphreys ont des conséquences qui dépassent l'étude de la science qu'il appelle « computationnelle » et sont riches de suggestions utiles à mon propos. Elles permettent en particulier de préciser la dimension que j'ai appelée « computationnelle » de l'activité théorique des agents.

⁷⁰Humphreys souligne qu'il existe des gabarits très puissants qui ne sont pas issus de théories. Cela rejoint les propos de Morgan et Morrison (1999), selon lesquelles les modèles sont des outils partiellement autonomes des théories.

⁷¹Il existe parfois une ambiguïté dans l'usage que Humphreys fait de ce terme. Tantôt les gabarits désignent les équations elles-mêmes, dans une forme particulière, tantôt la dimension syntaxique de ces équations, c'est-à-dire leur forme, considérée indépendamment de leur contenu. Il distingue ainsi parfois entre modèle computationnel et gabarit computationnel, ce dernier semblant caractériser la seule syntaxe du modèle. Mais souvent, Humphreys emploie « gabarit » pour parler du modèle lui-même, en tant qu'il est dans une forme particulière. Ce flottement est dû à la tension même qui caractérise la notion de gabarit, que je vais présenter à la section 3.2.3 (voir aussi la note 72).

système, elles soient en principe inter-déductibles et contiennent par conséquent les mêmes informations, elles ne permettent pas les mêmes calculs. Souvent même, un changement de gabarit – lié par exemple à un changement de coordonnées – permet de rendre soluble une équation qui, sous son ancienne forme – n'admet pas de solution analytique.

Un changement de coordonnées rectilignes en coordonnées polaires, du référentiel du laboratoire au référentiel du centre de masse, d'une représentation lagrangienne à une représentation hamiltonienne, d'une représentation qui n'autorise pas la séparation des variables à une représentation qui l'autorise – tout cela et plus encore fait la différence entre une représentation inutilisable computationnellement et une représentation utilisable computationnellement. (Humphreys, 2004, p. 97)

Une étude de l'application des représentations théoriques aux phénomènes exige, selon Paul Humphreys, que l'on donne toute son importance à ce type de différences. Pour cela, il convient de prêter attention au « morceau concret de syntaxe sur lequel le difficile travail d'application est mené » (pp. 99-100). La « syntaxe » dont il s'agit ici, et les opérations *concrètes* qu'elle permet d'effectuer, sont donc bien différentes des règles syntaxiques au nom desquelles l'approche abstraite des empiristes logiques leur permet de considérer les différentes formulations de la mécanique comme équivalentes.

L'approche de Paul Humphreys se présente non seulement comme une critique de la conception syntaxique, mais aussi de la conception sémantique des théories issue des travaux de Patrick Suppes. En effet, cette conception, tout en réhabilitant le rôle des mathématiques contre une conception strictement logiciste, restreint selon Humphreys « les mathématiques utilisées pour représenter les théories [...] à la théorie des ensembles » alors qu'il s'agit « au contraire de se concentrer de manière plus fine sur les théories effectivement utilisées dans les calculs » (Humphreys, 2004, p. 97). En restant à un niveau trop abstrait, les défenseurs de la conception sémantique « placent dans la même classe d'équivalence des formulations de théories qui sont, en pratique, radicalement différentes d'un point de vue computationnel » (p. 99).

3.2.2 Vers une nouvelle cartographie de l'activité théorique

En montrant qu'il convient de prêter attention à la forme des représentations, considérées indépendamment de leur contenu représentationnel, Paul Humphreys suggère que ces deux aspects sont orthogonaux : non seulement deux équations ayant le même contenu peuvent être dans des gabarits différents, mais encore on retrouve parfois les mêmes gabarits dans des domaines théoriques différents. Ainsi par exemple les techniques des paysages de fitness [*fitness landscapes*], mises au point dans le cadre de la biologie évolutionniste par Wright, Haldane et Fischer (voir Wright, 1932),

« peuvent-elles être appliquées à la biologie évolutionniste, à l'évolution de l'économie et aux stratégies de survie des corporations » (pp. 91-92). De même, les « automates cellulaires, inventés par von Neumann en guise d'outils pour représenter les automates auto-reproducteurs, ont été adaptés, élaborés et appliqués à des modèles multi-agents [*agent-based*] en sociologie, en économie et en anthropologie » (p. 92). Pour revenir à l'exemple de la mécanique classique, les principes variationnels qui sont au fondement des formulations analytiques se retrouvent également en optique ; quant à la forme des équations lagrangiennes d'une part et des équations hamiltoniennes d'autre part, on la retrouve dans la plupart des branches de la physique théorique.

En raison même du caractère trans-disciplinaire des gabarits, Humphreys suggère que l'attention aux gabarits invite à repenser le tracé même des frontières entre domaines scientifiques. En parallèle d'une individuation des théories par référence aux phénomènes qu'elles représentent, l'attention à la forme concrète des représentations et aux raisonnements et pratiques qui leur sont associés permet ainsi d'envisager, comme je l'avais suggéré au chapitre 2 (page 82), une reconfiguration de la carte de l'activité théorique. Cette nouvelle cartographie, qui ne recoupe pas celle des théories, est susceptible d'avoir, en raison même de l'importance grandissante des techniques informatiques dans la pratique scientifique, une réelle incidence sur les découpages disciplinaires. Les différents utilisateurs de paysages de fitness, qu'ils soient biologistes ou économistes, ont en effet des pratiques représentationnelles plus proches que deux biologistes n'utilisant pas les mêmes modèles théoriques pour étudier des phénomènes relevant pourtant d'un même domaine. S'il « est naturel, pour un scientifique expérimental, de prêter attention aux phénomènes concrets pour déterminer les frontières d'une science », « les choses sont différentes dans le domaine du théorique. Ici, la réorganisation est possible et arrive souvent » (p. 69). En faisant de la construction et de la manipulation de gabarits un élément important de l'étude de l'activité théorique, Paul Humphreys souligne ainsi l'aspect éminemment *pratique* de cette dernière, par opposition à la science expérimentale, que l'on considère pourtant traditionnellement comme la partie pratique de l'activité scientifique.

3.2.3 Représentation et calcul : une tension

L'insistance de Humphreys sur la dimension computationnelle de l'activité théorique, par opposition à une attention exclusive au contenu représentationnel *en principe* des équations, se traduit dans ce qu'il présente comme le slogan de son approche : « calcul, et non représentation » [*computation, not representation*] (p. 53). Cette opposition ne doit cependant pas masquer le fait que, en pratique, l'utilisation d'une équation dans un gabarit particulier est indissociable de son interprétation, comme

l'analyse des propos de Kuhn sur les cas exemplaires l'a mis en évidence. En effet, dans la mesure où l'analyse de l'activité théorique passe, selon Humphreys, par celle des « morceaux concrets de syntaxe » effectivement utilisés et non des « entités purifiées du philosophe » que sont les lois théoriques reconstruites par les approches formelles, elle ne peut se passer d'une attention à la tâche concrète en vue de laquelle les scientifiques utilisent tel ou tel gabarit. Il existe ainsi une tension, au sein de la définition même même des gabarits, entre l'idée d'une unité « purement syntaxique » indépendante d'un domaine particulier de phénomènes et celle d'un *outil concret* construit et transformé en vue de la représentation d'un phénomène particulier. C'est cette tension – reconnue et assumée comme telle par Humphreys – entre « l'usage trans-disciplinaire des gabarits formels et leur dépendance implicite au domaine étudié, qui en fait des unités d'analyse intéressantes. » (p. 71)

Paradoxalement, souligner l'importance de la syntaxe (au sens de Humphreys) des équations par opposition à leur contenu représentationnel implique du même coup de renoncer à la distinction entre syntaxe non interprétée et interprétation empirique, c'est-à-dire à la thèse de l'« *interprétation détachable* », selon laquelle « le gabarit est un pur morceau de formalisme sur lequel une interprétation peut être imposée » et selon laquelle « l'interprétation peut être retirée et remplacée par une autre interprétation ».

Adopter la conception de l'interprétation détachable revient à faire une grave erreur sur le statut épistémique des gabarits dans le travail scientifique. Si l'on prend les formes de gabarit de base pour des morceaux de syntaxe primitifs, on omet le fait que ces morceaux de syntaxe, dans la construction de ces gabarits, portent toujours en eux une interprétation visée et la justification qui lui est associée. Éplucher l'interprétation visée revient à retirer la raison qui justifie l'adoption de ces équations. (p. 80)⁷²

Si la syntaxe des équations utilisées compte, c'est en effet *du point de vue de leur utilisation en acte*, en vue d'une tâche particulière, et non pas du point de vue que Humphreys appelle la « perspective sans propriétaire » [*no-ownership perspective*] (p. 77) qui caractérise les reconstructions formelles des théories. Dans cette dernière perspective, comme on l'a vu, des équations de gabarit différent, comme les équations

⁷²Dans un article consacré aux modèles computationnels, Humphreys (2002, p. S10) écrit ainsi : « La question la plus intéressante est la suivante : à quelle étape dans le processus de construction et d'évaluation des modèles computationnels la connaissance propre au domaine étudié intervient-elle ? Je maintiens la distinction entre le gabarit qui n'est pas dépendant du domaine (au sens où il s'applique à une grande variété de domaines) et la connaissance nécessaire pour appliquer ce gabarit à un phénomène donné [...] L'ensemble des corrections est aussi toujours dépendant des phénomènes, et donc, malgré sa souplesse, le gabarit l'est aussi. Cela est en partie dû à l'inséparabilité du gabarit et de son interprétation. »

tions lagrangiennes et hamiltoniennes, sont équivalentes. Ce n'est donc que si l'on fait entrer en jeu l'utilisateur de ces équations et les calculs concrets qu'il peut faire à partir d'elles que la différence apparaît. Or, dès que l'on prend en compte un utilisateur particulier, et non un agent idéal, on doit prendre en compte le but dans lequel il utilise tel ou tel gabarit ; en conséquence, on ne peut pas ignorer l'interprétation sous-jacente au choix et à l'utilisation de ce gabarit. Les gabarits ne sont pas les axiomes non interprétés des théories ; ce ne sont pas des équations toutes faites « trouvées » dans les théories ; ce sont des outils théoriques *en train d'être utilisés*, et, par conséquent, toujours déjà interprétés. On retrouve ici l'idée kuhnnienne (présentée page 331) selon laquelle, même au niveau le plus général et abstrait, les équations utilisées par les membres d'une communauté scientifique sont interprétées, sans quoi elles ne pourraient pas être utilisées ; « une quantité considérable de connaissance du domaine étudié est nécessaire pour adapter correctement les gabarits aux cas actuels ».

Le choix de la fonction de force ou de la distribution de probabilité qui doit être utilisée pour transformer un gabarit théorique en un gabarit computationnel⁷³ exige une connaissance du domaine étudié – la « théorie » ne contient pas cette information et on a besoin de plus que de simples compétences inférentielles pour construire le gabarit computationnel. (p. 92)

Revenons au slogan cité précédemment (« calcul, et non représentation »). C'est au sens où le contenu représentationnel serait « détachable » des outils théoriques et devrait y être « attaché » par des règles d'interprétation ou par l'établissement d'un théorème de représentation entre structures que calcul (ou computation) et représentation s'opposent. En revanche, si l'on s'intéresse à la pratique concrète des scientifiques, il convient de prendre en compte le caractère indissociable de la manipulation et de l'interprétation, et d'étudier la manière dont la forme des représentations qu'ils utilisent détermine et est elle-même déterminée par la manière dont ils relient les différents éléments de ces représentations aux phénomènes qu'ils cherchent à expliquer. Cela permet, comme certains propos de Humphreys le suggèrent, de repenser à nouveaux frais le problème, présenté au chapitre 5, de ce que Nagel (1961) appelle le « statut cognitif » des théories.

⁷³Un gabarit théorique est une forme d'équation qui, en tant que telle, ne permet aucun calcul particulier, comme par exemple « $f = ma$ » ; le gabarit computationnel est la forme que prend cette équation une fois qu'on a spécifié le type de force en jeu dans le problème étudié (pendule simple, chute libre, etc.).

3.2.4 « Réalisme sélectif » (Humphreys, 2004, p. 82)

En raison même de la tension que je viens de souligner entre le caractère trans-disciplinaire des gabarits et le fait qu'ils sont toujours construits et utilisés dans un but précis, Paul Humphreys affirme qu'une analyse de l'activité théorique centrée sur les gabarits permet de comprendre que les agents n'adoptent ni ne rejettent les théories « en bloc », comme le suggère la perspective « sans propriétaire » – et comme le suggère aussi la thèse du holisme sémantique –, mais qu'ils font plutôt preuve d'un « réalisme sélectif » :

Le réalisme est souvent conçu comme une position à propos des théories, abstraction faite de l'attitude adoptée vis-à-vis d'elles. On considère qu'il concerne la question de la vérité ou de la fausseté des théories, ou celle de savoir si certains termes des théories bien confirmées font référence à des entités réelles. C'est ce que nous avons appelé [...] la perspective sans propriétaire. Mais les réalistes sont rarement si peu sélectifs, et prennent rarement autant de distance vis-à-vis du contenu de leurs théories. Le réalisme est ordinairement une attitude adoptée par un individu ou un groupe vis-à-vis de certaines parties d'une théorie, et les engagements réalistes ne sont pas directement lisibles dans [*read straight off*] la théorie elle-même. (p. 83)

Bien au contraire, nous dit Humphreys, les utilisateurs des théories sont souvent assez clairvoyants sur les parties de leurs modèles dont ils pensent qu'elles ont un référent actuel dans le monde empirique, et celles qui ne sont pas censées en avoir.⁷⁴ Selon les phénomènes qu'il sert à décrire, le même gabarit, choisi pour des raisons de facilité de calcul [*tractability*], n'est pas manipulé de la même manière : l'« ontologie sous-jacente » (p. 80) au système étudié guide son utilisation et son interprétation.

Ces considérations permettent ainsi de souligner l'interaction complexe, dans l'activité théorique, des impératifs liés au calcul et des processus interprétatifs, eux-mêmes guidés dans certains cas par les engagements théoriques des différents agents. Elles seront au cœur de l'analyse que je proposerai, au chapitre 9, des débats de certains généticiens des années 1920 sur la construction et l'interprétation des cartes génétiques. Pour clore ce chapitre, je voudrais à présent relier l'analyse que je viens de proposer de la dimension computationnelle de l'utilisation de représentations particulières dans la prédiction et l'explication des phénomènes avec la défense d'une sémantique inférentielle que j'ai esquissée dans les sections 1.3.2 et 1.3.3 du présent chapitre.

⁷⁴Pour une analyse de l'utilité des modèles faux – et reconnus comme tels par leurs utilisateurs –, voir l'article de William Wimsatt (1987).

3.3 Figures et formules : des unités conceptuelles

Dans l'introduction de ce chapitre, j'ai souligné la triple limitation de la notion de théorie : elle est à la fois trop large, parce que les théories ne sont jamais utilisées « en bloc », trop abstraite, parce qu'elles ne sont des outils de représentation et d'inférence qu'en tant qu'elles sont utilisées par des agents, et trop étroite, parce les représentations fournies par les théories tirent aussi leur signification du contexte – linguistique et scientifique – dans lequel elles sont utilisées, et des autres représentations auxquelles elles sont liées. Si l'analyse qui précède m'a permis de satisfaire les deux premières exigences – définir des unités locales et concrètes –, la dernière exigence a été momentanément laissée de côté. Dans cette courte section, je souhaite élargir les considérations qui précèdent dans une double direction, destinée à satisfaire cette troisième exigence. Le prochain chapitre vise à fournir les outils d'analyse permettant d'emprunter la voie que j'esquisse ici.

3.3.1 Les liens inférentiels entre représentations

En premier lieu, souligner l'importance de la forme particulière des représentations et des inférences qu'elles permettent aux agents de tirer, en pratique, pour prédire et expliquer le comportement de systèmes particuliers n'implique pas de réduire le rôle de ces représentations à la résolution de ce type de problèmes. Comme je l'ai affirmé en introduction, et comme je l'ai rappelé à la section 1.3.2 (page 312), les représentations tirent leur signification de la manière dont les agents les relient au monde empirique et *entre elles*. Autrement dit, affirmer que l'expression « $f = ma$ » tire *toute* sa signification de son application à des problèmes particuliers reviendrait à négliger le fait qu'elle fait aussi partie d'un ensemble de représentations avec lesquelles elle entretient des relations déductives.

L'analyse que je viens de mener permet cependant de préciser l'idée, suggérée à la section 1.3.2, selon laquelle la signification d'une représentation lui est donnée par son rôle inférentiel dans les raisonnements des agents. Elle nous enseigne en effet que l'étude de ce rôle inférentiel implique de prêter attention à la forme des représentations et à la manière dont les agents les manipulent ; cette forme est cruciale pour les inférences que les agents peuvent tirer au moyen de ces représentations, que ce soit lors de leur application à des problèmes particuliers, de leur utilisation dans l'apprentissage, ou du développement et de l'approfondissement d'une « théorie » (ou plus précisément d'une version), conçue comme un réseau de représentations liées entre elles. Cela permet de comprendre que deux équations « équivalentes » n'ont pas exactement la même signification si leur forme implique des manipulations différentes de la part des agents. Cela permet aussi de comprendre que la formule « $f = ma$ » n'a pas

exactement la même signification selon qu'on l'applique au mouvement de chute libre d'un corps ou à celui d'un pendule simple. Enfin, cela permet de comprendre qu'elle n'a pas non plus exactement la même signification selon qu'on l'apprend comme un principe fondamental ou qu'on la déduit d'un principe variationnel. Le prochain chapitre, au moyen de la notion de format, vise à approfondir cette perspective.

3.3.2 Différents types de représentation

Renoncer à analyser la signification des représentations théoriques sur le modèle de la référence linguistique invite en outre à élargir l'étude de l'activité théorique à des représentations non linguistiques. En effet, si l'on considère que la forme particulière des représentations joue un rôle important dans les raisonnements des agents, on doit prendre en compte *tous* les types de représentations utilisées dans la pratique scientifique, et non pas seulement les *formules*, si l'on entend par là les énoncés du langage naturel et les équations.⁷⁵ L'apprentissage, le développement et l'application des hypothèses théoriques se fait en effet très souvent au moyen de *figures*, c'est-à-dire de diagrammes, de graphes, de schémas, d'images picturales ou encore de modèles tridimensionnels.⁷⁶

L'importance des représentations spatiales – en particulier les diagrammes – dans la pratique scientifique a déjà été soulignée par certains philosophes des sciences (voir par exemple Baigrie, 1996; Griesemer et Wimsatt, 1989; Griesemer, 1990; Wimsatt, 1990; Woody, 2004; Perini, 2005; Giere, 2006), parfois inspirés par des travaux en sociologie des sciences (voir Lynch, 1988; Lynch et Woolgard, 1990). Souvent, une distinction nette est tracée entre l'utilisation de ce type de représentation et l'utilisation de représentations linguistiques : l'importance des représentations spatiales est soulignée par contraste avec celle des représentations linguistiques. Mon propos, dans le chapitre suivant, sera plutôt de fournir une analyse de l'interaction cognitive des agents avec les représentations qu'ils construisent et utilisent susceptible de rendre compte aussi bien de l'utilisation de formules que de figures. Autrement

⁷⁵L'attention exclusive à ce type de représentation est probablement due, chez Kuhn, d'une part à l'héritage empiriste logique et d'autre part au fait que, comme chez les empiristes logiques, la plupart de ses exemples sont tirés des sciences physiques. La perspective de Humphreys, qui vise à rendre compte de l'importance des ordinateurs dans la pratique scientifique, explique également que ses gabarits soient presque toujours des équations, ou en tous cas des schèmes de calculs descriptibles par des algorithmes ; les questions de calculabilité qui l'intéressent au premier chef ne représentent cependant qu'une partie des raisonnements et des manipulations concrètes mis en œuvre par les agents qui apprennent, développent et appliquent des théories.

⁷⁶Signalons d'ailleurs que, dans leur exposé de la sémantique des rôles conceptuels (voir ci-dessus page 313), Greenberg et Harman (2006) accordent une place importante à l'analyse de la signification des cartes, modèles et diagrammes (pp. 300-301).

dit, je me propose de concevoir tous ces types de représentation comme des unités *conceptuelles*.⁷⁷

Conclusion

L'analyse successive des notions de paradigme, de modèle, et enfin de cas exemplaire et de gabarit m'a permis de préciser la définition des unités d'analyse auxquelles il convient de prêter attention pour rendre compte de l'interaction cognitive des agents avec les représentations théoriques qu'ils utilisent. J'ai été progressivement conduite, de l'analyse d'unités globales, à concentrer mon attention sur des unités locales, les « morceaux concrets de théorie » que les agents construisent et manipulent au cours de l'apprentissage et de la pratique scientifiques. Qu'il s'agisse d'équations, de diagrammes ou de tout autre type de représentations, je propose de les considérer comme des unités isolables et relativement autonomes du cadre théorique duquel elles sont issues. Pour autant, leur rôle dans les raisonnements des agents ne peut pas être compris sans une étude de leur relation aux autres représentations qu'ils utilisent.

Conformément aux exigences énoncées en introduction, étudier ces « morceaux concrets » permet de rendre compte du fait que les agents, en pratique, ont rarement affaire à une théorie prise comme un « tout », mais plutôt à des représentations particulières, qui tirent leur signification de la manière dont ils les utilisent comme outils de représentation et d'inférence, qu'il s'agisse de les appliquer à la résolution de problèmes particuliers, ou d'explorer les conséquences déductives d'un principe. Un des buts du prochain chapitre est d'élaborer des outils d'analyse permettant de rendre compte de l'interaction entre la forme particulière de ces représentations et les capacités cognitives propres des agents.

⁷⁷L'étude des diagrammes de Weismann par Griesemer et Wimsatt (1989) est exemplaire d'une telle perspective. Ils se proposent ainsi d'étudier l'évolution conceptuelle en science, dans la lignée des travaux de Hull (1982), en prenant les diagrammes comme des « organismes conceptuels » et les théories desquelles ils sont issus comme des « environnements conceptuels », sachant que ces diagrammes peuvent être « exportés et modifiés dans différents contextes théoriques et didactiques – souvent dans des contextes où la théorie dont ils sont originellement issus fait l'objet d'un examen critique ou d'une attaque, et [où] ils sont [alors] employés au service de leur concurrente ». (p. 92) Sans que cela implique de prendre parti pour ou contre la perspective évolutionniste qu'ils défendent, mon approche doit beaucoup à la leur. Wimsatt et Griesemer suggèrent d'ailleurs que tous les types d'unités locales aisément isolables du contexte dans lequel on les trouve, comme des équations ou des diagrammes, sont susceptibles d'être de bonnes unités d'analyse de l'évolution conceptuelle.

Chapitre 7

Outils d'analyse des représentations théoriques



« Les diagrammes de Feynman dans la jungle amazonienne », tiré de R.D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problems*, New York, 1976 [1967], 120

Les deux chapitres qui précèdent m'ont permis d'approfondir la signification et de justifier l'adoption du principal présupposé qui guide ce travail : l'analyse du contenu des représentations théoriques ne peut se passer d'une attention aux agents qui utilisent ces représentations. En adoptant ce présupposé, j'ai cherché à montrer que seule une analyse minutieuse des processus par lesquels les agents construisent et manipulent des représentations particulières permet de comprendre l'articulation des deux fonctions qu'on attribue habituellement aux théories : une fonction de représentation des phénomènes empiriques, et une fonction d'outil inférentiel.

J'ai ainsi proposé, au chapitre précédent, de choisir des unités d'analyse locales et d'étudier l'interaction cognitive des agents avec ce que j'ai appelé les « morceaux de théorie » qu'ils utilisent. Ces « morceaux de théorie », dont les principaux exemples que j'ai pris jusqu'ici étaient des formules mathématiques (des équations), sont à la fois les expressions d'hypothèses théoriques et les représentations de phénomènes. C'est par leur moyen que l'on peut tirer des inférences à propos des phénomènes que l'on cherche à expliquer et à prédire.

Tout au long de ce travail, j'ai été conduite à souligner l'importance de la *forme* particulière de ces outils de représentation et d'inférence que les agents utilisent au cours de leur apprentissage et de leur pratique de la science. J'ai en effet cherché à montrer l'impossibilité d'une distinction nette entre la forme et le contenu des représentations théoriques.

Ce constat implique, comme je l'ai suggéré à la fin du chapitre précédent, d'élargir notre regard à des représentations *non exclusivement linguistiques*. En effet, une approche de l'activité théorique centrée sur la pratique ne peut ignorer l'usage très fréquent d'images, de diagrammes, de schémas et de graphes, aussi bien dans le contexte de l'enseignement que dans celui de la recherche. Prendre au sérieux la diversité des types de représentations utilisés dans l'apprentissage et la pratique scientifiques implique de renouveler nos outils d'analyse de ces représentations afin de rendre compte, dans le détail, des processus par lesquels les agents construisent, développent, et appliquent des hypothèses théoriques. C'est l'objet principal de ce chapitre.

Dans un premier temps (section 1), je présenterai la notion de *système symbolique*, empruntée à Nelson Goodman (1968/1976). Cette notion est destinée, chez Goodman, à caractériser différents *types de représentations* – linguistique, diagrammatique, pictural – en analysant la relation entre les propriétés perceptibles des représentations (les « marques » concrètes) et leur contenu. Elle m'offrira le cadre général de l'analyse des représentations que je propose dans la suite du chapitre. De plus, elle me permettra d'ébaucher la définition de certains types de représentations fréquemment utilisés dans la pratique scientifique, définition que je mobiliserai dans l'étude de cas de la troisième partie de ce travail (chapitres 8 et 9).

Ensuite (section 2), je centrerai mon analyse sur le problème particulier posé par les cas où deux représentations qui ont (au moins partiellement) *le même contenu* ne le présentent pas sous la même forme et, ce faisant, n'en facilitent pas également l'accès aux agents car elles ne leur permettent pas d'effectuer les mêmes inférences. De telles différences peuvent exister entre des représentations de types différents (comme par exemple une liste de données numériques et la présentation de ces mêmes données sous la forme d'un graphe), mais également entre deux représentations de même type (par exemple, deux graphes dans des systèmes de coordonnées différents). Suivant un usage courant, je propose, dans tous les cas de ce type, de parler de différence de *format*. Je chercherai à préciser la définition de cette notion intuitive de format au moyen du cadre d'analyse des représentations mis en place à la section 1.

Je montrerai enfin (section 3) que l'ébauche de définition proposée à la section 2 est insuffisante. Au moyen de deux exemples (les équations de la mécanique classique et les diagrammes de Feynman en électrodynamique quantique), je montrerai que l'analyse de la notion de format implique de façon irréductible la prise en compte des capacités cognitives propres ainsi que des compétences des agents et de leurs buts dans le contexte d'utilisation des représentations. En conséquence, on verra que la définition du contenu cognitivement accessible d'une représentation – la seule qui compte dans l'approche de l'activité théorique que je défends – implique la prise en compte d'agents particuliers dans une situation particulière. Je conclurai en esquisant quelques conséquences de cette analyse pour une définition de l'activité théorique.

1 Systèmes symboliques et types de représentations

Les outils traditionnels permettant d'étudier la relation entre la forme d'une représentation et son contenu, comme les notions de syntaxe et de sémantique, sont initialement destinées à l'étude des représentations linguistiques. En tant que telles, elles ne semblent pas pouvoir être appliquées à l'analyse de représentations comme les photographies, les dessins, les diagrammes ou les graphes. La construction et l'interprétation de telles représentations semblent en effet obéir à des règles très différentes de celles qui gouvernent les langages, aussi bien naturels que formels. On voit mal, par exemple, comment les outils destinés à étudier les règles de formation des énoncés à partir des mots, ainsi que les règles de formation des énoncés complexes à partir des énoncés plus simples pourraient s'appliquer à l'analyse d'une photographie. Afin de voir quels outils sont susceptibles de servir à l'étude des règles de construction et d'interprétation des représentations non linguistiques, prenons, pour commencer, quelques exemples simples.

Des traits noirs sur un papier blanc peuvent être considérés comme un ensemble de traces sans signification. Ils peuvent aussi, dans certains cas, représenter un objet. Cet « objet » peut être, selon les cas, toute sorte de choses. Ainsi par exemple, une « ligne noire ondulée sur fond blanc » peut aussi bien être la représentation graphique des battements d'un cœur (un électrocardiogramme) qu'un dessin de la ligne de crête du Mont Fujiyama (Goodman, 1968/1976, p. 229). L'image de la figure 14 joue ainsi sur la possibilité d'interpréter le même ensemble de traces visibles de différentes manières.

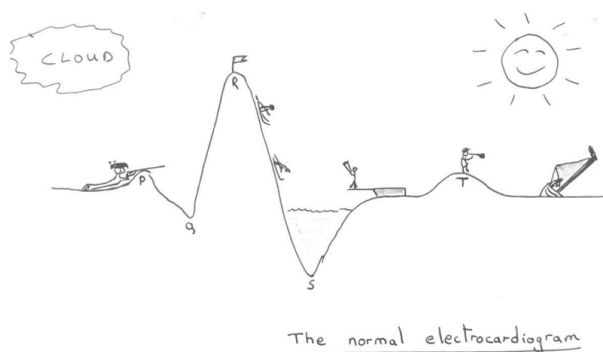


FIG. 14 – “L'électrocardiogramme normal.” In *Know more than your Doctor*, John Pollard, Gentle Pulse LTD.

De même, l'inscription $e=mc^2$ peut être lue comme une occurrence de la formule « $e = mc^2$ », au même titre que l'inscription $e = mc^2$, ou comme la photographie d'une inscription à la craie sur un tableau noir, comme c'est le cas dans la figure 15. Pour prendre un exemple similaire, les photographies de la figure 16 représentent,

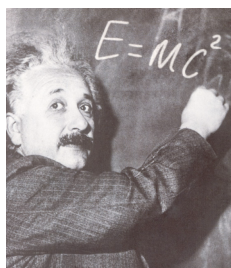


FIG. 15 – Albert Einstein.

entre autres, une statue. Cette statue (et non sa photographie) est une représentation tridimensionnelle d'Einstein. Confrontée au détail reproduit sur la droite, je peux y lire la formule « $e = mc^2$ », ou y voir une représentation non linguistique de cette inscription sur le cahier d'Einstein.



FIG. 16 – Mémorial d'Einstein à Washington.

Non seulement, selon les cas, ces différentes représentations n'ont pas le même objet – la ligne de crête d'une montagne, des battements de cœur, un être humain, une statue, un cahier rempli d'inscriptions, une relation entre des variables, etc. – mais encore, à supposer qu'elles aient le même objet (par exemple, des battements de cœur), elles ne nous apprennent pas la même chose à son propos : *elles ne contiennent pas exactement les mêmes informations*. De même qu'une ligne noire ondulée peut représenter des battements de cœur ou la ligne de crête du Mont Fujiyama selon la manière dont on l'interprète, le même objet peut être représenté par des représentations de types différents : « les montagnes peuvent être représentées diagrammatiquement et les battements de cœur picturalement » (Goodman, 1968/1976, p. 229). En revanche, il n'est pas certain qu'un diagramme et un dessin puisse nous apporter les mêmes informations à propos des battements d'un cœur. Pour prendre un autre exemple, comparons un énoncé sur Einstein et une photographie d'Einstein. Un énoncé, même assez bref, peut nous indiquer qu'Einstein était un grand physicien, né en 1879 à Ulm, inventeur de la théorie de la relativité, qui ne s'intéressait pas, pendant ses études primaires, à la biologie et aux sciences humaines. On voit mal comment une photographie pourrait nous apprendre tout cela. Inversement, un ensemble d'énoncés, même très long, pourrait difficilement nous renseigner aussi précisément qu'une photographie sur la forme de son visage. Comme l'indique la formule courante selon laquelle une image vaut parfois mieux que mille mots (et inversement), il semble bien que certains types de représentations soient plus adaptés que d'autres à l'expression de certains contenus.

Selon Goodman (1968/1976), l'objet de ces différentes représentations, ainsi que ce qu'elles nous disent à propos de cet objet – leur contenu informationnel – dépendent du *système symbolique* dans lequel elles sont interprétées. Cela signifie qu'il existe

un ensemble de règles de construction et d'interprétation en vertu desquelles ces « marques » représentent un certain objet et contiennent des informations à son propos. Au moyen de la notion de système symbolique, un des buts de Goodman est, comme on va le voir, de généraliser la notion de langage et, avec elle, celles de syntaxe et de sémantique, à l'analyse de tous les types de représentation. Il propose en effet de définir les règles de construction et d'interprétation de toutes les représentations (leur système symbolique) comme des règles syntaxico-sémantiques.

Après avoir, dans un premier temps (section 1.1), présenté la notion de système symbolique telle qu'elle est définie par Goodman, je me concentrerai sur l'usage des représentations à des fins épistémiques (section 1.2). Je préciserai alors, au moyen de la notion de système symbolique, ce qu'il faut entendre par « contenu informationnel d'une représentation ». Cela me permettra, dans un troisième temps (section 1.3), de définir certains types importants de représentations utilisées dans la pratique scientifique, et de montrer que cette diversité obéit à des impératifs authentiquement épistémiques (par opposition, par exemple, à des fins pédagogiques ou d'illustration).

1.1 La notion de système symbolique

Goodman définit le système symbolique auquel une représentation appartient (en vertu duquel elle a un certain contenu) par une analyse de ses propriétés perceptibles – les *marques* concrètes (visuelles, auditives, ou plus généralement accessibles aux sens)¹ – et des relations que ces marques entretiennent entre elles et avec ce qu'elles dénotent (leur objet). Ces relations sont définies au moyen des notions de syntaxe et de sémantique, que je présente dans un premier temps (section 1.1.1). Dans un deuxième temps (section 1.1.2), on verra que la définition des systèmes symboliques comme des ensembles de règles syntaxico-sémantiques permet à Goodman de distinguer différents types de représentations (pictural, linguistique, diagrammatique) par la définition de certaines propriétés caractéristiques des systèmes symboliques dans lesquels elles fonctionnent. Autrement dit, selon Goodman, ce qui différencie, par exemple, la photographie d'un chien de sa description linguistique, réside dans certaines propriétés structurelles (syntaxico-sémantiques) des systèmes symboliques (pictural, linguistique) dans lesquels ces représentations sont interprétées.

1.1.1 Syntaxe et sémantique des systèmes symboliques

La syntaxe d'un système symbolique – que Goodman appelle son « schème symbolique » – consiste en un ensemble de *caractères* et en leurs règles de combinaison.

¹Goodman se restreint à l'analyse des marques ; on pourrait cependant élargir son analyse à tout objet matériel, y compris des objets tridimensionnels.

Les caractères sont des classes de *marques*. Le schème symbolique du français écrit, par exemple, comprend les lettres de l'alphabet romain, ainsi que des caractères composés suivant certaines règles au moyen de ces caractères élémentaires que sont les lettres. Le mot « oculiste » est un caractère de ce schème symbolique – et une représentation dans le système symbolique qu'est le français écrit – et comprend toutes ses occurrences écrites (inscriptions²).

Les propriétés perceptibles d'une marque ne sont pas toutes pertinentes pour définir son identité syntaxique, c'est-à-dire pour déterminer de quel caractère elle est une instance. Un changement syntaxique implique un changement dans l'aspect de la marque, mais l'inverse n'est pas toujours vrai : dans certains schèmes symboliques, comme celui du français écrit, on peut modifier dans une certaine mesure l'aspect d'une marque sans que cela modifie son identité syntaxique. Ainsi, par exemple, un changement dans la police ou dans la couleur de l'encre utilisée ne modifie pas la marque³ a au point d'en faire l'instance d'un autre caractère (une autre lettre) : a est du même type syntaxique que a et a. Ce sont toutes des instances de la lettre « a ». En revanche, la marque d n'est pas une instance du caractère « a », mais du caractère « d ». Elle n'est pas du même type syntaxique que les marques précédentes. De même, les inscriptions oculiste, oculiste, et oculiste sont des occurrences du même caractère du français écrit, le mot « oculiste ». En revanche, la marque ocullste, pour infime que puisse sembler la modification, n'est pas une instance du mot « oculiste ».

Reprenons l'exemple de la marque $e=mc^2$. Si on la considère comme une représentation linguistique (en l'occurrence, une formule mathématique), elle est syntaxiquement identique à la marque $e = mc^2$. Toutes deux sont des occurrences de l'expression « $e = mc^2$ ». En revanche, si l'on considère la marque $e=mc^2$ comme la photographie d'une inscription faite à la craie sur un tableau, les règles syntaxiques du système dans lequel elle fonctionne (grossièrement, celui des images photographiques) ne sont pas les mêmes ; dans ce cas, une modification de la couleur ou de l'épaisseur du trait sont des modifications syntaxiques. Elles affectent l'identité syntaxique – et, comme on le verra, sémantique – de cette représentation : l'inscription $E=mc^2$ n'est pas une instance du même caractère que l'inscription $e=mc^2$.

À la suite de John Kulvicki (2003, 2006b), j'appellerai « propriétés syntaxique-

²Les inscriptions sont les marques qui sont les instances d'un caractère (Goodman, 1968/1976, p. 131).

³J'adopte la convention suivante : j'utilise des guillemets pour mentionner les caractères (en l'occurrence, les lettres et les mots), c'est-à-dire les classes de marques, et je présente les marques elles-mêmes dans un encadré.

ment pertinentes » les aspects perceptibles des marques en vertu desquelles elles sont les occurrences d'un caractère, et donc en vertu desquelles elles sont d'un certain type syntaxique.⁴ Ainsi, dans le schème symbolique du français écrit dans son usage standard, la couleur des marques n'est pas une propriété syntaxiquement pertinente, alors qu'elle l'est dans le schème symbolique d'une photographie en couleurs. En revanche, dans aucun des deux cas, la texture du papier n'est (en général) une propriété syntaxiquement pertinente.⁵

La sémantique d'un système symbolique assigne un domaine de référence – ou « classe de concordance » – à son schème et impose des règles gouvernant la manière dont les différents caractères du schème dénotent⁶ les éléments du domaine.⁷ Dans les systèmes non ambigus⁸, deux instances d'un même caractère ont le même référent. Autrement dit, l'identité syntaxique implique l'identité sémantique. En revanche, l'inverse n'est pas toujours vrai. En effet, des inscriptions ayant le même référent peuvent appartenir à des caractères différents : dans les langages naturels, les cas de synonymie sont un exemple d'identité sémantique de deux caractères différents. Les marques oculiste et ophtalmologue sont des occurrences de caractères différents. Pourtant, en français, ils ont la même classe de concordance.

Caractériser les propriétés syntaxiques et sémantiques des représentations permet de comparer entre eux différents systèmes, sur la base des propriétés qui leur sont communes. Cela permet, comme on va le voir, de dégager différents types de systèmes (et, par conséquent, de représentations).

⁴Comme le remarque Kulvicki (2006b, p. 16, p. 29), parler des *propriétés* des marques responsables de leur appartenance à un type syntaxique est incompatible avec la position nominaliste de Goodman, qui en dit peu sur ce point, se contentant d'évoquer les « aspects constitutifs d'un caractère » [*character constitutive aspects*]. Comme Kulvicki, j'adopte le cadre goodmanien sans que cela implique un engagement en faveur du nominalisme.


⁵Pour plus de détails sur cette notion et sur la caractérisation complète de l'ensemble des propriétés syntaxiquement pertinentes d'un schème, voir (Kulvicki, 2006b, pp. 30-31).



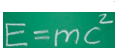


⁶Goodman distingue entre « référence » et « dénotation ». Une représentation peut selon lui référer sans dénoter : par exemple, « une peinture de couleur grise *ne dénote pas* la couleur grise mais *est dénotée* par le prédicat “gris”. » (p. 52) Ce faisant, la peinture « exprime » le gris, et Goodman considère que l'expression est un mode de référence. Il réserve le terme « dénotation » aux cas que l'on appelle, en général, indifféremment « référence » et « dénotation ». Dans la mesure où je ne m'intéresserai pas aux cas de ce que Goodman appelle « expression », je suivrai moi-même l'usage courant et ne distinguerai pas entre dénotation et référence.

⁷Goodman (p. 143) précise qu'il faut entendre « dénotation » au sens large : on doit par exemple pouvoir considérer un système dont le schème symbolique est un ensemble de lettres et de mots écrits et où la classe de concordance est l'ensemble des sons qui correspondent à leur prononciation.

⁸Voir (Goodman, 1968/1976, p. 147).

1.1.2 Quelques propriétés importantes des différents types de systèmes

Reprenons l'exemple de la marque . C'est en vertu d'un certain système symbolique qu'elle fonctionne comme représentation et qu'elle acquiert un contenu. Ainsi, c'est le système symbolique dans lequel elle est interprétée qui détermine, selon Goodman, son objet (une relation mathématique ou une portion de tableau noir sur laquelle figure une inscription à la craie). Ce système, en fixant les règles syntaxico-sémantiques qui en permettent la lecture, décide en effet du type de représentation dont il s'agit. En l'occurrence, il peut s'agir d'une représentation de type *linguistique* ou d'une représentation de type *pictural*. Les représentations picturales sont celles dont on dit couramment qu'elles « ressemblent » à leur référent ; les photographies, ainsi que les peintures et les dessins « réalistes » sont des exemples typiques de représentations picturales.⁹

Comme je l'ai brièvement montré dans la section précédente, les règles en vertu desquelles la marque  fonctionne comme une représentation linguistique ou comme une représentation picturale sont en effet différentes. On a vu, par exemple, que la syntaxe du schème symbolique en vertu duquel cette marque est une formule mathématique est beaucoup moins *sensible* à certains changements de forme que celle d'un schème pictural. Pour le dire autrement, alors que  et  correspondent exactement au même caractère (et ont exactement le même contenu¹⁰) en vertu d'un système linguistique, ce n'est pas du tout le cas si on les interprète comme des représentations picturales : dans un cas, l'objet représenté est un tableau noir sur lequel figure une inscription d'une certaine forme ; dans l'autre, l'objet représenté est un tableau vert sur lequel figure une inscription assez différente (ce n'est qu'en vertu d'un système linguistique que l'on peut affirmer que des inscriptions de forme aussi différente que  et  sont les instances d'un même caractère).

Un des buts visés par Goodman dans *Langages de l'art* est de caractériser différents types de représentation – linguistique, pictural, diagrammatique, etc. – en se

⁹Signalons que Goodman réserve le terme « représentation » à ce que j'appelle ici les représentations « picturales ». Il appelle « symbole » la classe plus large de ce que j'ai appelé jusqu'ici « représentations », y incluant les langages, les images picturales, les diagrammes, les graphes, etc. (voir Goodman, 1968/1976, p. 4, note 1) Goodman appelle par conséquent « systèmes représentationnels » [*representational systems*] le type particulier de systèmes symboliques qui caractérise les représentations picturales. Je suivrai l'usage courant qui consiste à appeler « représentation » tout type d'objet tenant lieu d'autre chose. Les systèmes que Goodman appelle « représentationnels » seront donc appelés « systèmes picturaux ».

¹⁰Comme je l'ai dit (page 354), l'identité syntaxique n'implique pas toujours l'identité sémantique, en raison des cas de synonymie. Cependant, dans cet exemple, les deux marques sont identiques syntaxiquement et sémantiquement.

référant *exclusivement* aux propriétés structurelles (syntaxico-sémantiques) des systèmes symboliques dans lesquels elles fonctionnent. Cela signifie, par exemple, qu'il entend caractériser les systèmes picturaux sans faire référence à la notion de ressemblance, au moyen de laquelle j'en ai précédemment donné une définition intuitive¹¹ : comme je l'ai dit, ce qui différencie, selon lui, la photographie d'un chien d'une description linguistique réside dans les propriétés syntaxico-sémantiques des systèmes symboliques dans lesquels ces représentations fonctionnent, et non dans la similitude de l'expérience visuelle de la photographie et du chien lui-même. Cette affirmation concernant la picturalité a été très discutée (voir par exemple Files, 1996; Perini, 2004)¹² ; je ne prétends, dans ce travail, ni l'approfondir ni la critiquer.¹³ En revanche, ce sont les outils que Goodman élabore dans ce but que je souhaite reprendre à mon compte. Ils me seront en effet utiles pour distinguer entre quelques types importants de représentations utilisées dans la pratique scientifique ; je me propose de les présenter dans la suite de la présente section.

Goodman commence par définir les systèmes qu'il appelle « notationnels » (voir Goodman, 1968/1976, chap. 4). Ces systèmes, dont un des exemples paradigmatiques chez Goodman est celui qui régit les partitions musicales, obéissent à des réquisits syntaxiques et sémantiques stricts ; c'est par la manière dont ces différents réquisits sont satisfaits ou violés que Goodman définit les autres systèmes. Comme on va le voir, les systèmes linguistiques satisfont les réquisits syntaxiques des systèmes notationnels – ils ont un schème notationnel – mais violent pour la plupart leurs réquisits sémantiques. Les systèmes diagrammatiques et picturaux, à leur tour, ne satisfont généralement aucun des réquisits des systèmes notationnels ; ils se distinguent entre eux par d'autres propriétés syntaxiques.

Voyons d'abord en quoi consistent les *réquisits syntaxiques* des systèmes notationnels, qui définissent donc les schèmes notationnels. Une condition nécessaire pour qu'un schème soit notationnel est que les différentes instances de chacun de ses caractères soient interchangeable : cela implique que si deux marques sont des occurrences du même caractère, il est impossible que l'une des deux soit l'occurrence d'un caractère sans que l'autre le soit aussi. C'est ce que Goodman appelle l'*indifférence de*

¹¹ Voir (Goodman, 1968/1976, chap. 1) pour sa critique de la notion de ressemblance.

¹² Pour des analyses de la picturalité centrées, à la différence de celle de Goodman, sur l'expérience perceptuelle, voir (Wollheim, 1974, 1980, 1993, 1998; Gombrich, 1960, 1982; Hopkins, 1998; Schier, 1986; Peacocke, 1987; Budd, 1993). L'approche de Dominic Lopes (1996) est mixte, au sens où elle inclut des critères structurels comme ceux de Goodman tout en accordant une place à l'expérience perceptuelle.

¹³ Pour une analyse approfondie de l'entreprise goodmanienne consistant à caractériser le pictural en termes structurels, voir les travaux de John Kulvicki (2003, 2006a,b), qui se présentent comme un prolongement critique de ce projet.

caractère (Goodman, 1968/1976, pp. 131-133). Une conséquence de cette première condition est le réquisit de *disjonction des caractères* (Goodman, 1968/1976, pp. 133-135), selon lequel aucune marque ne peut appartenir à plus d'un caractère. En effet, si une marque appartenait à deux caractères, alors, soit ces deux caractères seraient identiques, soit la condition d'indifférence de caractère serait violée, puisque certaines marques seraient des instances de l'un des deux mais pas de l'autre. Une deuxième conséquence de la condition d'indifférence de caractère est le réquisit de *différentiation finie des caractères* (Goodman, 1968/1976, pp. 135-136) : cela signifie qu'il est toujours possible, au moins en principe, de déterminer à quel caractère une marque appartient. Les systèmes linguistiques, comme je l'ai dit, satisfont ces réquisits syntaxiques.¹⁴ Il est en effet toujours possible, sauf cas anormaux, de savoir de quelle lettre, par exemple, une marque donnée est une instance. En revanche, les systèmes linguistiques ne satisfont pas, en général, les réquisits sémantiques des systèmes notationnels, que je ne fais que mentionner rapidement, dans la mesure où ils n'entrent pas en jeu dans la distinction entre systèmes linguistiques, diagrammatiques et picturaux, qui sont ceux qui m'intéressent au premier chef.


Ces *réquisits sémantiques* sont les suivants. Le premier est l'absence d'ambiguïté, selon laquelle aucun caractère ne doit pouvoir être relié à plus d'une classe de concordance. Le deuxième est la disjonction sémantique, selon laquelle la classe de concordance d'un caractère ne doit pas avoir des éléments communs avec la classe de concordance d'un autre caractère. Enfin, le troisième réquisit sémantique est la différenciation finie sémantique, qui exige qu'il soit toujours possible de dire à quel caractère un élément de la classe de concordance correspond.

Voyons à présent comment ces notions, et plus particulièrement celles qui concernent les règles syntaxiques, permettent de distinguer entre deux ensembles de représentations, les représentations linguistiques d'une part, et les représentations diagrammatiques et picturales de l'autre. Alors que les langages, comme je l'ai dit, satisfont les réquisits syntaxiques des systèmes notationnels, les graphes et les images picturales n'obéissent pas au réquisit de différenciation finie des caractères. Leur syntaxe, à la différence de celle des langages, est *inarticulée*. La violation du réquisit de dif-

¹⁴Le fait que leur syntaxe soit ainsi articulée (qu'elle obéisse au réquisit de différenciation finie des caractères) est indispensable, entre autres, à leur compositionnalité, c'est-à-dire au fait que la signification des caractères complexes est déterminée par la signification et l'arrangement des caractères atomiques.

férentiation finie des caractères définit la *densité* d'un schème¹⁵ : un « schème est syntaxiquement dense s'il contient une infinité de caractères ordonnés de telle façon qu'entre chacun d'entre eux, il y en ait toujours un autre. » (Goodman, 1968/1976, p. 136) :

les systèmes non linguistiques diffèrent des langages, le dessin de la description, le pictural¹⁶ du verbal, les peintures des poèmes, en premier lieu par un manque de différenciation – en fait, par la densité (et en conséquence une totale absence d'articulation) – du système symbolique. Rien n'est intrinsèquement une représentation [picturale] ; le statut de représentation [picturale] est relative à un système symbolique. (Goodman, 1968/1976, p. 226)

Prenons, encore une fois, l'exemple de la marque . La *sensibilité* (voir ci-dessus, page 355) du schème symbolique en vertu duquel elle fonctionne comme représentation picturale est une conséquence de la densité syntaxique : dans la mesure où, entre chaque caractère de ce schème, il en existe toujours un autre, la moindre modification de l'aspect de cette marque en fait l'instance d'un autre caractère (dans le cas contraire, il y aurait une solution de continuité entre les caractères du schème).

De même que la densité syntaxique est une violation de la différenciation finie de caractères, la densité sémantique est une violation de la différenciation finie sémantique. Un système est sémantiquement dense quand les différents éléments de sa classe de concordance sont ordonnés de telle façon qu'entre chacun d'eux, il en existe toujours un autre. Densités syntaxique et sémantique sont indépendantes ; pour des exemples des différents cas, je renvoie à (Goodman, 1968/1976, pp. 57-59) et, au prix de quelques imprécisions sans effet sur mon propos, je n'entrerai pas dans ces détails. Parmi les représentations auxquelles je m'intéresse ici, celles dont la syntaxe est dense sont également denses sémantiquement.¹⁷ Une représentation picturale, comme par exemple une photographie, mais aussi un graphe représentant l'évolution temporelle d'une variable, sont denses (syntaxiquement et sémantiquement). À l'ensemble dense des caractères qui en constituent le schème correspond un ensemble dense de ré-

¹⁵Plus précisément, c'est la violation *complète* du réquisit de différenciation finie qui définit la densité. L'absence de densité ne garantit pas, à elle seule, la différenciation finie (voir Goodman, 1968/1976, p. 137). Cela est sans effet sur l'usage que je ferai, dans la suite de ce chapitre, des outils goodmaniens.

¹⁶Goodman parle ici de « représentationnel ».

¹⁷Les représentations dont la syntaxe et la sémantique sont denses sont les représentations que Goodman appelle « analogiques ». La distinction entre analogique et digital correspond, intuitivement, à celle entre continu et discret. Cette distinction a fait l'objet de nombreuses définitions différentes et de plusieurs débats, notamment en philosophie de l'esprit, où elle prend un sens assez différent. Voir (Lewis, 1971; Dretske, 1981; Haugeland, 1981; Block, 1981; Peacocke, 1986, 1987; Blachowicz, 1997; Katz, 2008).

férents. C'est la raison pour laquelle tout changement dans la forme des marques implique une modification syntaxique et une modification sémantique : le contenu de la représentation qui résulte d'une telle modification n'est pas le même que celui de la représentation de départ. Une modification, même infime, dans la forme de la ligne noire ondulée représentant le Mont Fujiyama en change le contenu : elle ne me dit plus exactement la même chose sur la ligne de crête du Mont Fujiyama. De même, si je modifiais, même très légèrement, les inscriptions gravées sur le cahier d'Einstein (figure 16), je modifierais le contenu de la représentation, au sens où l'écriture manuscrite d'Einstein ne serait plus représentée exactement de la même manière.

La notion de densité permet de distinguer, comme on vient de le voir, entre les formules (représentations linguistiques) et certaines figures comme les images picturales et les graphes ; elle ne permet cependant pas de distinguer ces dernières les unes des autres. La condition supplémentaire qui permet de distinguer entre les représentations picturales et les graphes ou certains diagrammes, est la condition de *saturation relative*. Il s'agit, encore une fois, d'une condition syntaxique. La saturation d'une représentation se caractérise par la quantité de propriétés syntaxiquement pertinentes.¹⁸

Les seuls aspects pertinents du diagramme sont les ordonnées et les abscisses de chaque point où le centre de la ligne passe. L'épaisseur de la ligne, sa couleur et son intensité, la taille absolue du diagramme, etc., n'ont pas d'importance ; ces aspects n'ont pas d'importance pour déterminer si une copie de la représentation appartient au même caractère du schème diagrammatique. En revanche, pour l'esquisse, ce n'est pas vrai. Une modification dans l'épaisseur de la ligne, sa couleur, le contraste avec l'arrière-plan, sa taille, et même les qualités du papier – aucun de ces aspects ne peut être écarté ou ignoré. Quoique les schèmes diagrammatiques et picturaux soient identiques du point de vue de la non articulation, certains aspects constitutifs du schème pictural sont contingents dans le schème diagrammatique ; les représentations du schème pictural sont relativement *saturées*. (Goodman, 1968/1976, p. 230)

Cela signifie que, bien que le graphe soit dense (la moindre modification de la hauteur d'un point est une modification syntaxique et sémantique du graphe), certaines de ses propriétés perceptibles ne sont pas pertinentes : il a un moins grand nombre de propriétés syntaxiquement pertinentes. C'est ce qui fait que la ligne ondulée, si elle est interprétée comme la représentation des battements d'un cœur, n'est pas picturale ou, plus précisément, qu'elle est « moins picturale ». En effet, la saturation – et donc la picturalité – est, telle que Goodman la définit, affaire de degrés : « un schème dense est plus diagrammatique [moins pictural] qu'un autre si les aspects constitutifs des

¹⁸ Goodman parle de la quantité des « aspects constitutifs des caractères ». Voir ci-dessus, note 4.

caractères dans le premier sont inclus dans ceux du second. » (Goodman, 1968/1976, p. 230)

Récapitulons brièvement les distinctions permises par les notions que je viens d'exposer. Les systèmes diagrammatiques et picturaux se différencient des systèmes linguistiques par la densité de leur syntaxe (et de leur sémantique). Ils se distinguent entre eux par leur saturation relative. Comme on va le voir, en permettant l'analyse de la relation entre la structure (syntaxico-sémantique) d'une représentation et son contenu, ces outils nous offrent le moyen d'examiner les raisons pour lesquelles différents types de systèmes symboliques sont plus ou moins adaptés à l'expression de différents types d'informations, et donc à l'atteinte de certains buts épistémiques (pour reprendre la formule consacrée, les raisons pour lesquelles une image vaut parfois mieux que mille mots). Avant d'examiner ainsi certains types de représentation courants dans la pratique scientifique, je me propose, dans la section suivante, d'approfondir l'analyse de ce que j'appelle le « contenu informationnel » d'une représentation.

1.2 L'usage épistémique des représentations et leur contenu informationnel

Les représentations remplissent des fonctions diverses¹⁹ : épistémiques, esthétiques, ou encore mnémoniques. Je me concentrerai exclusivement, dans ce qui suit, sur la fonction épistémique des représentations, c'est-à-dire sur leur usage par des agents dans le but d'acquérir des connaissances à propos de la chose représentée. Cette « chose » pouvant être un objet matériel, certaines propriétés de cet objet, un processus, l'évolution d'une valeur, etc., je parlerai, de façon neutre, de la « cible » de la représentation. Toute représentation, quel que soit son usage dominant, est susceptible de *contenir un certain nombre d'informations* à propos de sa cible.

Dans la mesure où ce qui m'intéresse ici est l'usage épistémique des représentations, je me concentrerai sur leur contenu informationnel, à l'exclusion, par exemple, de ce que Goodman appelle leur contenu « expressif ». Sans entrer dans l'analyse des difficultés soulevées par la notion d'information²⁰, je me contenterai de définir une information comme ce qui peut faire l'objet d'une attitude propositionnelle (que la proposition en question soit exprimée linguistiquement ou non). Le contenu informationnel d'une représentation peut être initialement défini comme *l'ensemble des informations qu'un agent maîtrisant le système symbolique dans lequel elle fonctionne*

¹⁹ Je ne parle ici que des fonctions que les représentations remplissent *en tant que* représentations : l'utilisation d'une toile de maître pour cacher un trou sur un mur ne fait pas partie de ces fonctions.

²⁰ Voir (Dretske, 1981; Cohen, 2004; Floridi, 2004, 2009).

peut en principe en tirer. Cette définition appelle plusieurs éclaircissements.

1.2.1 Les deux « étapes » de l'utilisation épistémique d'une représentation

On peut décrire l'utilisation d'une représentation en vue d'obtenir des informations à propos de sa cible comme consistant en deux étapes. Ces étapes sont souvent simultanées, mais il convient ici de les distinguer. Prenons par exemple le portrait du duc de Wellington présenté ci-dessous (figure 17). Supposons que je cherche, par son moyen, à obtenir des informations sur l'aspect physique du duc de Wellington. Ma maîtrise des systèmes picturaux me permet d'apprendre, par exemple, qu'il avait la peau blanche, la bouche assez fine et les cheveux châtons. Pour que mon entreprise épistémique – celle qui consiste à chercher des informations à propos de Wellington dans ce tableau – soit couronnée de succès, il faut, entre autres, que la croyance qui en résulte soit une croyance vraie ; il faut que les informations extraites du tableau soient correctement interprétées comme étant à *propos de sa cible*. Autrement dit, si je conclus de mon examen que le duc de Wellington avait la peau particulièrement blanche alors que le peintre a (volontairement ou non) exagéré la blancheur de sa peau, mon entreprise échoue. Il me faut donc, pour pouvoir utiliser avec succès cette représentation, disposer d'un certain nombre de renseignements à propos de la source (être certaine que le tableau a bien été peint à partir du duc de Wellington), de la manière dont les données concernant cette source ont été recueillies et de la précision avec laquelle elles sont présentées dans le tableau.



FIG. 17 – Portrait du duc de Wellington, par Sir Thomas Lawrence.

Considérons à présent les graphes de la figure 18. Si j'interprète la courbe des températures comme indiquant qu'il fait exactement 25°C tout au long du mois de juillet à Paris, alors que cette valeur représente la moyenne des températures prises chaque matin à 8 heures, mon entreprise épistémique échoue.²¹ Il en va de même si je crois que ces graphes concernent Madrid alors que les mesures ont été prises à Paris, ou encore si je crois qu'ils concernent l'évolution annuelle de la température à Paris en 2008 alors qu'ils représentent une moyenne obtenue au cours des vingt dernières années. Une erreur d'interprétation de ce type peut être due à plusieurs

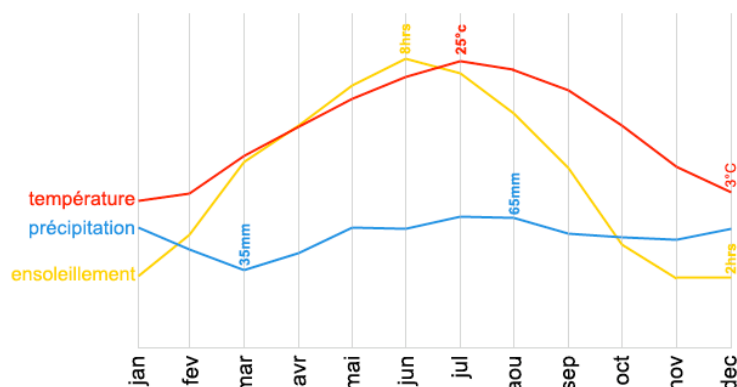


FIG. 18 – Graphe représentant l'évolution annuelle de la température, des précipitations et de l'ensoleillement à Paris.

causes, comme par exemple une méconnaissance de la manière dont les données ont été rassemblées ou des procédures d'approximation préalables à la construction de la représentation. Je peux également être victime d'une tromperie volontaire de la part de l'auteur de la représentation, qui me fait croire qu'il a collecté des données à tel endroit alors qu'il les a en fait obtenues ailleurs, ou même inventées. Toutes ces erreurs d'interprétation concernent l'adoption de croyances à propos d'une certaine cible ; l'adoption de ces croyances repose sur *une inférence, à partir de certaines propriétés de la représentation, à des propriétés de sa cible*. Je propose, pour plus de clarté, d'appeler cette inférence l'étape n°2 (car elle est seconde, sinon chronologiquement, du moins logiquement) de l'utilisation d'une représentation à des fins épistémiques.

Avant d'inférer – correctement ou non – des propriétés du graphe aux propriétés de sa cible, je dois *savoir lire le graphe lui-même* afin d'en extraire des informations. C'est là l'autre étape de l'utilisation épistémique d'une représentation (en fait, logiquement première, que j'appellerai donc l'étape n°1). Pour un graphe donné, le

²¹Sauf bien sûr si j'ai conscience du caractère approximatif de cette proposition, et si elle me suffit pour atteindre le but que je vise.

système dans lequel il fonctionne détermine lesquelles de ses propriétés perceptibles sont syntaxiquement pertinentes et comment elles doivent être interprétées *au sein du graphe lui-même*. Extraire des informations du graphe – avant même de les interpréter comme portant sur une cible particulière – implique de maîtriser non seulement les règles syntaxiques du système, mais également ses règles sémantiques, lesquelles ne se réduisent pas à la connaissance de sa cible effective. Autrement dit, le graphe contient des informations *tout court*, que ces informations puissent être correctement interprétées comme portant sur une certaine cible ou non. Quand un enseignant trace un graphe au tableau afin d'apprendre à ses élèves à lire ce type de représentations, il n'a pas besoin de supposer que ce graphe représente l'évolution d'une quantité réelle. Néanmoins, on peut raisonnablement affirmer que le graphe « dit » que la quantité représentée augmente ou diminue au cours du temps. De même, un portrait de Julien Sorel dit des choses à propos de Julien Sorel, alors même que ce dernier n'existe pas. Il contient bien cette information, que l'agent se trompe ou non en l'attribuant à une certaine cible, et qu'il existe ou non une telle cible.²²

Maîtriser un système symbolique consiste à savoir extraire les informations d'une représentation fonctionnant dans ce système : cela exige d'être en mesure d'en identifier les propriétés syntaxiquement pertinentes et de savoir les interpréter correctement au sein même de la représentation, c'est-à-dire *avant* même d'en conclure que la cible de cette représentation a telle ou telle propriété. La distinction de ces deux étapes sera de première importance pour la suite de mon analyse ; je m'y intéresserai en effet presque exclusivement à la seconde étape exposée (l'étape n°1, chronologiquement première), celle pour laquelle la maîtrise du système symbolique est cruciale. Il convient cependant de nuancer cette distinction. Il existe en effet de nombreux cas ambigus, et la distinction entre ce qui relève de la maîtrise du système symbolique permettant d'extraire des informations et ce qui relève de leur interprétation à *propos* de la cible de la représentation n'est pas nette.

Pour comprendre cela, reprenons l'exemple du duc de Wellington. J'ai donné comme exemple d'échec dans l'étape n°2 le fait de conclure que la peau de Wellington est particulièrement blanche alors que le peintre a exagéré cette blancheur. Si l'on admet que l'un des codes picturaux de la peinture anglaise de la fin du dix-huitième

²²La question de savoir si l'on peut effectivement parler de contenu informationnel quand il n'existe pas de référent ne va pas de soi ; l'examiner me conduirait inévitablement à aborder la difficile question des fictions. La solution de Nelson Goodman à ce problème consiste à dire qu'un portrait représentant Julien Sorel sous les traits d'un homme brun est une image appartenant à une certaine catégorie (celle des portraits-d'homme-brun) mais ne représentant (ne dénotant) rien. Ainsi, un graphe représentant l'augmentation d'une valeur sans référent serait un graphe-d'augmentation-de-valeur ne représentant rien.

siècle consiste à exagérer la blancheur de la peau des aristocrates, alors mon erreur n'est pas due à un manque de renseignements sur la manière dont les données ont été collectées, mais bien à une mauvaise maîtrise du système symbolique dans lequel ce portrait fonctionne. Si j'étais plus experte, j'aurais été en mesure de rectifier l'information, au sein même du tableau ; plus précisément, je n'aurais pas du tout tiré cette information du tableau. Un bon exemple de cette difficulté nous est fourni par les représentations métaphoriques ou caricaturales.²³ Ainsi, la caricature du duc de



FIG. 19 – Caricature du duc de Wellington par Auguste Louvrier de Lajolais.

Wellington (figure 19) nous le présente recroquevillé sur lui-même et doté d'un très long nez tout rouge. Dois-je tirer ces informations de l'image, que j'interprète alors littéralement, pour ensuite en décrypter le caractère métaphorique, ou bien ma maîtrise des systèmes symboliques caricaturaux me permet-elle directement de tirer de cette image des informations du type « Wellington était un poltron alcoolique » ?²⁴

Ces considérations valent aussi pour le cas des représentations utilisées dans la pratique scientifique. Ainsi, prenons l'exemple d'une image obtenue au moyen d'un microscope. Il est difficile de déterminer si l'élimination du « bruit » lié à la manipulation (par exemple, la présence d'une poussière dans l'appareil) fait partie de la

²³Goodman y consacre une analyse approfondie, que je ne présenterai pas ici (voir Goodman, 1968/1976, pp. 85-90).

²⁴Cet exemple révèle une difficulté supplémentaire que je n'examinerai pas : ces informations peuvent être interprétées comme des informations à propos de Wellington, mais également comme des informations sur le caricaturiste lui-même (sur ses croyances à propos de Wellington). Plus généralement, toute représentation contient des informations sur le point de vue depuis lequel cette représentation a été construite. J'ignorerai autant que possible ce problème, qui m'éloignerait trop de mon propos.

maîtrise du système symbolique de ce type d'image (étape n°1), ou bien si cette compétence intervient au moment de l'interprétation de l'image comme étant à propos de l'objet étudié au microscope (étape n°2).

1.2.2 La détermination contextuelle du contenu des représentations


La notion de « maîtrise d'un système » recèle un autre type de difficultés. Certaines informations contenues dans une représentation y figurent de manière explicite²⁵, comme le fait qu'il fait 25°C à Paris en juillet dans le graphe de la figure 18. D'autres sont accessibles *via* certaines inférences, comme par exemple le fait qu'il fait 22°C de plus en juillet qu'en décembre. Selon les types de système, savoir tirer ces inférences exige une expertise plus ou moins grande : l'interprétation des propriétés de la représentation peut consister en l'application de règles de transformation plus ou moins complexes qui permettent à l'agent de tirer des inférences à propos de son contenu.

Selon que l'on considère ou non que la maîtrise de ces règles fait partie de la maîtrise du système symbolique, on jugera ou non que les informations obtenues par l'application des règles font partie de son contenu (en vertu de la définition du contenu informationnel donnée ci-dessus, page 360). Or, la réponse à cette question dépend fortement du type de représentations et du contexte de leur utilisation. Par exemple, il serait déraisonnable d'affirmer que la maîtrise du système symbolique dans lequel les équations différentielles fonctionnent n'inclut pas la maîtrise de tout un ensemble de règles de calcul ; si je suis simplement capable, face à des équations newtoniennes, d'en tirer les informations qui y sont explicitement présentées (la valeur des variables représentant la force et la masse, et la relation de proportionnalité entre force et masse), on ne peut pas dire que j'en maîtrise vraiment le système. Si l'on s'en tient à ce qui est explicitement représenté par ces équations, leur contenu est bien pauvre. En revanche, rien ne permet d'affirmer que la maîtrise des règles de conversion entre degrés Celsius et degrés Fahrenheit fasse partie de la maîtrise du système symbolique du graphe de la figure 18. En conséquence, la question de savoir si ce graphe contient l'information selon laquelle il fait 77°F à Paris en juillet (information que l'on peut obtenir au moyen de ces règles) n'a pas de réponse déterminée et dépend du contexte d'utilisation de ce graphe.

Pour finir, il est difficile de savoir quelle quantité de connaissance d'arrière-plan (et non pas seulement quelles règles de transformation) entre dans la détermination

²⁵La notion d'explicité est loin d'aller de soi. À ce sujet, voir par exemple (Kirsh, 1991). Je la prends ici en un sens intuitif, et je reviendrai sur le sens précis auquel je l'entends plus loin, avec la notion d'extractibilité immédiate (section 2.2.2).

du contenu d'une représentation. Ici aussi, cela semble dépendre du contexte. Réduire le contenu d'une représentation à ce que John Haugeland (1991) appelle son contenu « squelettique » [*bare-boned*]²⁶ ne rend pas compte du fait que, en pratique, de très nombreuses connaissances d'arrière-plan sont mobilisées pour la lecture des représentations les plus communes, jouant le rôle de prémisses implicites dans les raisonnements que l'on mène au moyen de ces représentations.

Le contenu d'une représentation dépend donc fortement du contexte dans lequel elle est utilisée, et cela en plusieurs sens. D'abord parce que le contexte détermine, comme on l'a vu, les connaissances et les compétences supposées partagées par les utilisateurs. Ensuite parce que c'est le contexte dans lequel elle apparaît et certaines indications annexes comme, par exemple, la légende qui accompagne certaines images, qui suggèrent le système dans lequel elle doit être interprétée. Pour reprendre l'exemple de Goodman, c'est le contexte dans lequel elle apparaît qui m'indique si la « ligne noire ondulée » est un électrocardiogramme ou un dessin du Mont Fujiyama. De même, c'est le contexte dans lequel la marque  apparaît qui m'invite ou non à faire abstraction du fond gris et de l'aspect blanc et irrégulier des marques, qui ne sont pas des propriétés syntaxiquement pertinentes dans le système symbolique du langage mathématique mais le deviennent dans le système de la photographie.

Pour toutes ces raisons, il est impossible de déterminer le contenu d'une représentation sans prendre en compte le contexte dans lequel elle apparaît. Cependant, une fois le contexte fixé et, par conséquent, une fois que l'on a précisé l'ensemble des hypothèses implicites et les règles du système, le contenu informationnel d'une repré-

²⁶ John Haugeland (1991) distingue entre le contenu « squelettique » [*bare-boned*] d'une représentation et son contenu « charnu » [*flesh out*]. Le contenu squelettique est le contenu « strict », qui n'est augmenté ou obtenu par la médiation d'aucune autre connaissance. Pour reprendre l'exemple de Haugeland, un rapport de police indiquant que, dans la chambre où un corps inanimé a été trouvé, les tiroirs étaient retournés, les vêtements en désordre et déchirés, les meubles abîmés et des verres brisés au sol, et le portefeuille du mort retrouvé au sol, sans argent liquide ni carte de crédit, ne dit pas, à strictement parler, qu'il y a eu une violente bagarre et que la victime a été volée. Ces dernières propositions font partie du contenu charnu de la description, mais pas de son contenu squelettique. Comme le remarque Haugeland, cette distinction ne recoupe pas la distinction entre contenu explicite et contenu implicite – à supposer que cette distinction ait elle-même un sens précis – : ce qui est implicite peut être inféré de ce qui est explicite, sans ajout d'hypothèse ou de connaissance supplémentaire (mais au moyen de règles d'inférence). Pour reprendre l'exemple du graphe, il contient implicitement l'information selon laquelle il ne fait pas 30°C à Paris en juillet, puisqu'elle peut être déduite de celle selon laquelle il faisait 25°C, sans hypothèse supplémentaire. En revanche, affirmer que la victime a été volée à partir de la description mentionnée ci-dessus implique une certaine connaissance du monde. Il est cependant difficile, comme le reconnaît Haugeland, d'établir nettement quelle part de la connaissance commune peut être considérée comme une connaissance implicite.

sensation est entièrement déterminé par les marques qui la constituent. Cela signifie que la question de savoir s'il existe des agents capables d'appliquer ces règles et la manière dont ils les appliquent (les inférences particulières qu'ils mettent en œuvre) n'entrent pas dans la définition du contenu informationnel de la représentation. En cela, il est distinct de ce que j'appellerai, à la section 2 du présent chapitre, son contenu *cognitivement accessible*.

Les représentations utilisées dans la pratique scientifique se caractérisent entre autres par les compétences et les connaissances d'arrière-plan que leur utilisation requiert de la part des agents. Après avoir précisé ce que j'entends par « représentation théorique », j'emploierai les outils goodmaniens pour ébaucher une définition de certains types de représentations théoriques particulièrement importants pour mon propos.

1.3 Vers une typologie des représentations théoriques

Des équations différentielles aux images obtenues par résonance magnétique, en passant par les diagrammes moléculaires, les graphes et les cartes génétiques, les représentations utilisées dans la pratique scientifique sont de types très divers. Certaines représentent des objets concrets ; c'est le cas, par exemple, de la plupart des images obtenues au moyen de techniques d'imagerie comme la microscopie, ou des schémas tirés de telles images. D'autres représentent des processus ou des mécanismes, comme par un exemple le diagramme de la synthèse d'une protéine. D'autres encore, comme les graphes ou les équations, représentent la relation entre des variables. Chacune de ces représentations peut être utilisée pour étudier un objet ou un processus *particuliers* (une zone particulière du cerveau d'un être humain, le mouvement de la Lune autour de la Terre) ou pour acquérir (ou transmettre) des connaissances à propos d'un *type* d'objets ou de processus (la structure des cellules humaines en général, la loi générale d'évolution du pendule).

Ce qui distingue les représentations que j'appelle « *théoriques* » des autres représentations à usage épistémique (par exemple, une carte du relief de la France ou le graphe de la figure 18) est qu'elles ne sont pas seulement la présentation de *données*, mais aussi l'expression d'*hypothèses*. Prenons par exemple l'image d'un caryotype (figure 20). Cette représentation peut être utilisée pour obtenir des informations à propos de l'individu particulier d'où la cellule est issue, à des fins, par exemple, de diagnostic prénatal ; elle peut aussi servir – comme c'est le cas, par exemple, dans un manuel – à l'acquisition de connaissances générales sur la cellule et les chromosomes. De plus, c'est, entre autres, la possession de ces connaissances théoriques (sur les chromosomes humains en général) qui permet d'extraire les informations

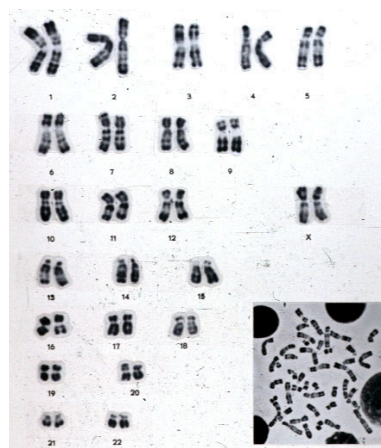


FIG. 20 – Caryotype (<http://medidacte.timone.univ-mrs.fr>).

contenues dans cette image et, le cas échéant, d'en tirer des inférences à propos de l'individu concerné. Ces connaissances théoriques sous-tendent, dans une certaine mesure, la maîtrise même des règles du système. Pour parvenir à isoler correctement les différents chromosomes et pour les organiser sous la forme d'un caryotype (pour « transformer » l'image en bas à droite en un caryotype proprement dit) afin d'en tirer des informations sur le sexe et les éventuelles anomalies génétiques de l'individu concerné, il me faut posséder un certain nombre de connaissances génétiques et cytologiques. La maîtrise du système symbolique dans lequel ce type de représentation fonctionne est donc en partie constituée de ces connaissances théoriques. C'est parce que je suis en mesure de l'interpréter *comme un caryotype en général* que je peux lire ce caryotype particulier.

Comme cet exemple, ainsi que ceux des équations différentielles et des images obtenues au microscope (mentionnés dans la section précédente), le montrent, les représentations théoriques requièrent de leurs utilisateurs des compétences particulières (par exemple, la maîtrise du calcul différentiel, la capacité à faire abstraction du « bruit » dans une image obtenue au microscope afin d'isoler ses propriétés syntaxiquement pertinentes), et des connaissances théoriques qui sous-tendent la maîtrise de ces règles et jouent le rôle de prémisses implicites dans les raisonnements. Ces compétences et ces connaissances, contrairement à celles qui sont requises, par exemple, pour la lecture d'une image picturale, ne sont partagées que par un petit nombre d'individus (les experts). Cette caractérisation ne suffit pas, certes, à tracer une frontière nette entre les représentations théoriques et celles qui ne le sont pas ; affirmer que les connaissances requises pour la lecture d'une représentation théorique sont des connaissances théoriques ne fait visiblement que repousser la question. Pour autant,

il n'est pas indispensable, pour mon propos, de tracer une telle frontière. On peut admettre que, de la lecture d'une carte routière ou d'un graphe comme celui de la figure 18 à celle d'un diagramme représentant l'interaction de particules microscopiques dans l'espace-temps, la différence est de degré. Dans les deux cas, la maîtrise de certaines règles syntaxico-sémantiques et la possession de certaines connaissances générales sont nécessaires.²⁷ Cette différence de degré est cependant considérable : les hypothèses théoriques contenues dans les représentations scientifiques leur procurent une puissance prédictive et explicative que les représentations communes n'ont pas.

Dans ce qui suit, je ne prétends aucunement proposer une typologie stricte et encore moins exhaustive des représentations utilisées dans la pratique scientifique. Il en existe une telle diversité que cette tâche paraît assez irréalisable ; en outre, ces représentations sont d'une telle sophistication que leur classement reposerait nécessairement sur des choix conduisant à négliger plusieurs aspects importants de leur contenu et de leur usage. Pour finir, les types que l'on peut dégager se trouvent rarement, sinon jamais, à l'état « pur » : un des aspects essentiels de l'utilisation de représentations diverses dans la pratique scientifique réside justement, comme on va le voir, dans leur interaction et leur enrichissement mutuel. Je ne prendrai donc que quelques exemples très simples. Ils me permettront, d'une part, de montrer comment l'adoption de la notion de système symbolique éclaire le fait que différents types de représentation, en raison même des propriétés syntaxico-sémantiques du système symbolique dans lequel elles fonctionnent, sont mieux adaptés que d'autres à l'expression de certains contenus, et donc à l'atteinte de certains buts épistémiques. D'autre part, ces quelques exemples me fourniront l'occasion de fixer le vocabulaire que j'utiliserai dans la suite de ce travail (en particulier dans les chapitres 8 et 9).

Dans un premier temps (section 1.3.1), je distingue les représentations spatiales des représentations linguistiques. Ensuite, je présente successivement trois grandes catégories de représentations spatiales : les représentations picturales (section 1.3.2), les représentations schématiques (section 1.3.3), et enfin les représentations diagrammatiques (section 1.3.4).

1.3.1 Représentations linguistiques et représentations spatiales

Bien que mon parti pris méthodologique consiste à étudier tous les types de représentations utilisés dans la pratique scientifique sans attribuer aux représentations linguistiques un statut essentiellement différent, j'admets une différence initiale entre

²⁷Pour une défense de l'idée selon les représentations scientifiques ne sont pas essentiellement différentes, en tant que représentations, des autres représentations externes, voir (Callender et Cohen, 2006).

deux genres de représentation : les représentations linguistiques ou *formules* et les représentations spatiales ou *figures*. Les formules incluent les expressions en langage mathématique comme les équations, les énoncés en langage naturel (augmenté de termes techniques), ainsi que les listes (par exemple, une liste de données numériques correspondant à un relevé de température). Les figures peuvent être unidimensionnelles ou bidimensionnelles (on pourrait élargir l'analyse aux représentations tridimensionnelles, mais je me concentrerai pour plus de simplicité sur les figures bidimensionnelles) et incluent les images picturales, les schémas, les diagrammes, les cartes, les graphes, mais également les tableaux de données.

Afin de distinguer entre ces deux classes de représentations, il convient de définir un critère qui n'est pas fourni par les outils goodmaniens. En effet, comme le souligne Laura Perini (2004), la distinction entre les systèmes dont la syntaxe est articulée (qui satisfont le réquisit de différenciation finie des caractères) et ceux dont la syntaxe est inarticulée (dont le schème est dense) ne permet pas de rendre compte de la différence entre les langages et certaines figures comme, par exemple, les diagrammes moléculaires (figure 21). Un tel diagramme a une syntaxe articulée : les différentes

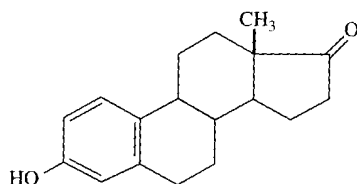


FIG. 21 – Diagramme d'une molécule d'oestrone (tiré de Perini, 2004).

marques qui le constituent peuvent être identifiées comme les instances d'un caractère particulier. Les lignes y dénotent des liens chimiques entre atomes, et les lettres y dénotent des types d'atomes. La longueur précise des lignes n'est pas une propriété syntaxiquement pertinente : sa modification n'implique pas un changement de caractère. De même, les lettres tenant lieu de types d'atomes ne sont pas plus sensibles à un changement de forme que ne le sont les lettres d'un énoncé en langage naturel. Du point de vue goodmanien, ce diagramme ne semble pas exclu de la catégorie des systèmes linguistiques ; en tout cas, il n'entre pas dans celle des images picturales et de ce que Goodman appelle « diagrammes ».

Pourtant, il semble bien qu'il y ait un sens intuitif de la notion de spatialité (ou de figure) qui permette de classer un diagramme de ce type dans la catégorie large des représentations spatiales (ou figures), qui partagent toutes, par delà leur diversité, une propriété dont les représentations linguistiques (ou formules) sont dépourvues.

Pour définir cette propriété, qui permettrait donc de distinguer entre, d'une part, *toutes* les figures, qu'elles soient denses ou non, et, d'autre part, les formules, il semble que les critères purement syntaxiques ne soient pas suffisant ; il est nécessaire d'établir un critère sémantique. Reprenons le diagramme de la figure 21. Comme je l'ai dit, la longueur des lignes n'y importe pas. Cependant, l'*emplacement* des lignes relativement aux lettres représentant les atomes et leur emplacement relatif (entre elles) ont une signification. Une ligne dont l'extrémité est à proximité d'une lettre représente un lien chimique avec l'atome désigné par la lettre ; l'intersection de deux lignes dénote un atome de carbone. Ainsi, certaines *relations spatiales* entre les caractères dont cette figure est composée sont *par elles-mêmes* significatives.

Ce n'est pas le cas d'un système linguistique : l'identité des caractères complexes (mots et phrases) y est déterminée uniquement par la *séquence* des caractères élémentaires (lettres). Mais les relations spatiales en tant que telles ne sont pas significatives. Ainsi, je propose de définir les figures, par opposition aux formules, comme des représentations au sein desquelles *certaines relations spatiales dénotent en tant que telles certaines propriétés de leur référent*.²⁸ Cette distinction semble pouvoir rendre compte de l'idée selon laquelle la lecture d'un énoncé linguistique suit l'ordre séquentiel des lettres et des mots, alors que celle d'une figure ne suit pas, en général, d'ordre déterminé (en tout cas, elle ne suit pas un ordre séquentiel).

À la différence de certains philosophes qui cherchent à rendre compte de l'importance des représentations non linguistiques dans la pratique scientifique, comme par exemple Laura Perini (2004), à qui l'analyse qui précède est partiellement empruntée, je parlerai de représentations *spatiales* et non *visuelles*, pour plusieurs raisons. D'abord parce que dans la perspective d'analyse des représentations en termes de systèmes symboliques, les langages écrits, par exemple, sont aussi visuels que les images picturales. Ensuite, parce que l'analyse que j'en fournis n'accorde pas de statut particulier à l'expérience visuelle en elle-même ; rien ne s'oppose *a priori* à ce que cette analyse s'applique à l'équivalent tactile de ces images. Enfin, parce que la notion même de *visualisation* soulève plusieurs problèmes. Selon le type de représentation dont il s'agit, elle ne semble pas avoir la même signification : des expressions comme « visualiser un objet grossi au microscope », « visualiser un processus non spatial ou la relation entre deux variables », et enfin « visualiser une formule mathématique au moyen de sa représentation graphique » n'emploient manifestement pas ce terme dans le même sens. Un des buts de l'ébauche de typologie qui suit est de préciser ces différents sens. J'y présente successivement trois grandes catégories de représenta-

²⁸En un sens très lâche de la notion d'isomorphisme, on peut dire que les représentations spatiales sont toujours partiellement isomorphes à leurs référents. Voir (Lee, 1999) pour une analyse de la représentation picturale au moyen de la notion d'isomorphisme.

tions spatiales : les représentations picturales, les schémas, et enfin les diagrammes.²⁹ Comme on le verra, ces catégories sont elles-mêmes assez hétérogènes ; en outre, la frontière qui les sépare n'est pas toujours parfaitement déterminée.

1.3.2 Les représentations picturales

Comme je l'ai dit précédemment, les représentations picturales sont celles dont on dit, intuitivement, qu'elles « ressemblent » à leur référent. Dans la pratique scientifique d'aujourd'hui, les représentations picturales par excellence sont toutes celles qui sont obtenues au moyen de techniques d'imagerie plus ou moins sophistiquées, comme la microscopie optique et électronique, la résonance magnétique, la radiographie, ou encore les techniques utilisées en astrophysique. Autrefois, le dessin et la gravure étaient des représentations picturales courantes dans certains domaines scientifiques (par exemple, l'anatomie ou la botanique).

Indépendamment des difficultés, soulignées par Goodman (1968/1976), que pose la notion de ressemblance pour la définition des représentations picturales courantes comme les photographies et les dessins et peintures dits « réalistes », cette notion semble particulièrement difficile à appliquer dans le cas des représentations scientifiques. En effet, la plupart du temps, la cible de ces images n'est pas elle-même *visible* : il s'agit le plus souvent d'objets très petits ou très éloignés, ou dont les propriétés mises en évidence par la technique ne sont pas détectables à la lumière blanche, mais au moyen de rayons X, de rayons Gamma, d'ultra-sons, ou encore d'électrons, pour ne citer que quelques exemples. Pour cette raison, entre autres, il semble particulièrement souhaitable, dans le cas des représentations scientifiques, de chercher à définir la picturalité d'une représentation sans faire appel à l'expérience perceptuelle de son utilisateur.

Les représentations picturales sont typiquement des représentations de « choses », c'est-à-dire d'entités concrètes, dont elles permettent de mettre en évidence certaines propriétés en respectant la *structure spatiale* de ces entités (et la distribution spatiale de ces propriétés). Ainsi, par exemple, l'image de la figure 22 est obtenue par une technique de microscopie électronique en transmission. L'image est le résultat des effets d'interaction entre des électrons et l'échantillon étudié ; le rayonnement X provoqué par le faisceau électronique permet en outre d'obtenir des informations concernant la composition chimique des tissus. Cette technique ne consiste donc pas seulement à

²⁹ Je ne prétends pas ici rendre compte de l'usage courant de ces mots : les termes de schéma et de diagramme sont parfois utilisés comme synonymes, et il est difficile de savoir ce qui les distingue dans l'usage habituel. Mon but, dans ce qui suit, est simplement de souligner des différences importantes entre certains types de représentation, et de fixer un vocabulaire qui est lui-même relativement arbitraire.

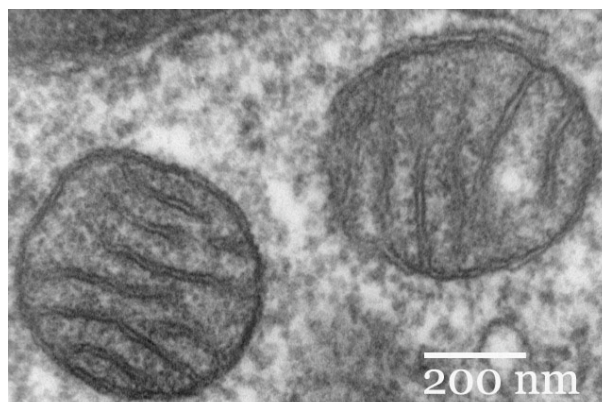


FIG. 22 — Micrographe électronique de cellules pulmonaires permettant d'observer des mitochondries (Louisa Howard, Dartmouth Electron Microscopy Facility, <http://remf.dartmouth.edu/imagesindex.html>).

« grossir » l'image d'un objet que l'on pourrait observer à la lumière blanche s'il était macroscopique ; les propriétés mises en évidence (comme les propriétés chimiques) ne sont pas visibles, au sens où ce que l'on observe est leur interaction avec les électrons et non avec de la lumière blanche. Cependant, la répartition spatiale des marques qui représentent ces propriétés non visibles respecte la structure spatiale de la cellule.

Ces considérations me permettent d'énoncer un premier élément de la définition des systèmes symboliques picturaux (qui leur est commun, comme on le verra, avec les systèmes schématiques, mais qui les différencie des systèmes diagrammatiques) : les systèmes symboliques picturaux se caractérisent en premier lieu par le fait que les relations spatiales au sein d'une représentation fonctionnant dans un tel système *représentent des relations spatiales*. On peut, bien sûr, observer des distorsions : les distances relatives sont plus ou moins bien respectées, mais ces représentations *respectent les relations topologiques*, sinon les relations métriques, de leur cible. Intuitivement, cela correspond à l'idée selon laquelle deux représentations picturales du même objet, quelles que soient les propriétés qu'elles permettent de visualiser, peuvent être projetées l'une sur l'autre. Le résultat de cette projection serait une troisième représentation picturale, qui permettrait entre autres de mettre en relation les propriétés de la cible présentées dans les deux représentations initiales.

Les deux autres critères permettant de caractériser les représentations que j'appelle « picturales » sont ceux de densité et de saturation, tels que Goodman les définit (voir ci-dessus, section 1.1.2). Dans la mesure où une représentation peut être partiellement dense et où la saturation est affaire de degrés, la picturalité l'est aussi : les images les plus picturales sont entièrement denses et très saturées. Cela signifie

que la moindre modification de leur aspect en change la signification, et qu'elles ont un très grand nombre de propriétés syntaxiquement pertinentes – intuitivement, cela correspond à l'idée selon laquelle elles sont très « remplies » : la plupart de leurs propriétés perceptibles sont signifiantes. C'est la satisfaction relative de ces deux critères qui distingue les représentations picturales des représentations schématiques, que je présenterai dans la prochaine section.

Les images picturales obtenues par des techniques d'imagerie comme celle présentée ci-dessus subissent la plupart du temps de nombreuses modifications et ajouts, destinés d'une part à éliminer le « bruit » lié à la manipulation des appareils et les aspects non pertinents qui brouillent la lecture, et d'autre part à en souligner les aspects pertinents : on peut en colorier différentes zones (comme dans la figure 23), en grossir certains traits, ajouter des indications linguistiques ou des symboles comme des flèches, destinés à *aider la lecture* de l'image. Ces différentes manipulations sont

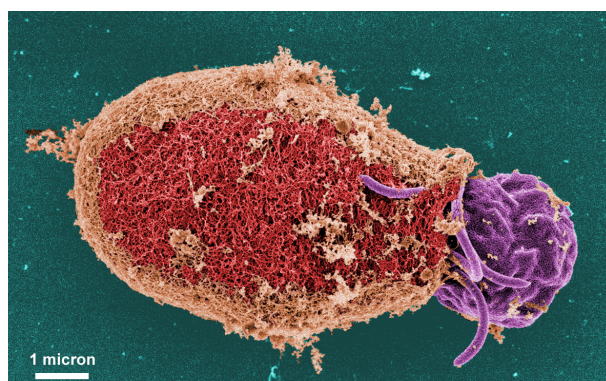


FIG. 23 – Micrographe électronique colorisé du protozoaire *Giardia* (Dr. Stan Erlandsen, Centers for Disease Control and Prevention, Public Health Image Library, [http ://phil.cdc.gov/phil/home.asp](http://phil.cdc.gov/phil/home.asp)).

intrinsèquement liées au statut *théorique* de ces images : elles contribuent à expliciter les hypothèses théoriques qui en permettent la lecture.

Ces opérations consistant à mettre en valeur les propriétés syntaxiquement pertinentes de l'image afin de la rendre plus lisible et d'expliciter certaines hypothèses théoriques s'accompagnent souvent de la suppression d'autres aspects. Par exemple, en coloriant certaines zones, on rend moins visibles certains détails jugés non pertinents. Cela contribue souvent à diminuer la saturation de l'image, et à introduire, en son sein, des éléments qui ne sont pas, eux-mêmes, denses. Je propose d'appeler ces différentes opérations d'abstraction des opérations de *schématisation* ; leur résultat est une image moins picturale, et plus *schématique*.

1.3.3 Les représentations schématiques

Tout comme les représentations picturales, les représentations que je propose d'appeler « schématiques » respectent les relations topologiques de leur cible. Elles peuvent donc être projetées sur une représentation picturale de la même cible. Cependant, à la différence des représentations les plus picturales, elles peuvent n'être que partiellement denses, et elles sont moins saturées. La figure 24 présente ainsi deux images : celle du haut est la représentation picturale, obtenue au microscope électronique, d'une mitochondrie ; celle du bas est un schéma tiré de cette image. Comme

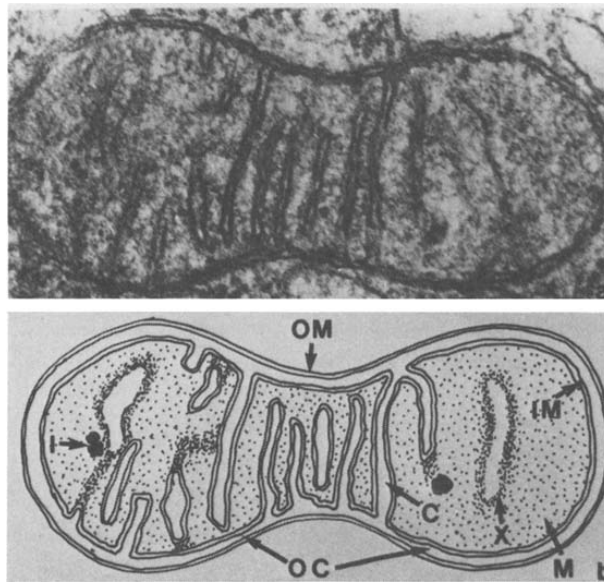


FIG. 24 – Mitochondrie du pancréas d'une souris. Micrographe électronique (en haut) et schéma (en bas). OM et IM : membranes internes et externes ; I : granule intramitochondrial ; C : crête ; X : section oblique d'une crête ; OC : compartiment externe ; M : matrice. Les deux images sont empruntées par Lynch (1988) à (Threadgold, 1976, p. 321).

le montre Lynch (1988), le processus qui permet d'obtenir le schéma à partir du micrographe ne consiste pas seulement en la simplification et la sélection de certains aspects de l'image, mais aussi en un certain nombre d'opérations qui contribuent à « identifier dans le spécimen particulier des propriétés "universelles" » (Lynch, 1988, p. 205). Ces opérations consistent à rendre le schéma plus uniforme, en prolongeant certaines lignes qui apparaissent interrompues dans le micrographe et en augmentant le contraste entre les différentes parties de la mitochondrie. Elles consistent aussi à accorder certaines propriétés visibles de l'image avec « l'identité que l'on assigne à certaines entités » (Lynch, 1988, p. 209) : ainsi, par exemple, la frontière entre

l'intérieur et l'extérieur de la mitochondrie est beaucoup plus nette dans le schéma. Enfin, le schéma « définit » les différentes parties qu'il contribue à rendre visibles au moyen de labels linguistiques et de flèches (voir légende). Ce faisant, le schéma ajoute un grand nombre d'informations qui ne sont pas directement visibles dans le micrographe et nous enseigne à *les y voir*. En ce sens, l'utilisation du schéma apprend à lire le micrographe comme une représentation théorique.

Dans cette figure, le schéma, comme le micrographe, respectent les relations topologiques de la cible. Cependant, en raison de l'introduction de flèches et de symboles linguistiques, le schéma n'est que partiellement dense : interrompre la ligne qui marque la frontière de la mitochondrie modifierait l'identité syntaxique de l'image (cette ligne est donc dense) ; en revanche, grossir les flèches ou les lettres, voire modifier légèrement la forme des frontières de la mitochondrie ne ferait pas du résultat un schéma différent.³⁰ En outre, le schéma est beaucoup moins saturé que le micrographe : chaque détail perceptible de ce dernier est syntaxiquement pertinent, alors que, dans le schéma, l'épaisseur du trait marquant le contour de la mitochondrie et la couleur du fond, par exemple, ne sont pas des propriétés syntaxiquement pertinentes. Les modifier ne changerait pas l'identité syntaxique ni, par conséquent, l'identité sémantique du schéma (son contenu).

Un schéma est également moins saturé qu'un dessin, par exemple, dans lequel les hachures représentant les ombres contribuent à en faire une représentation picturale. Notons d'ailleurs que certains articles de vulgarisation introduisent parfois une dimension picturale dans des représentations schématiques, en ajoutant, par exemple, des ombres ou des reflets sur les entités représentées (par exemple, dans la figure 25, sur des cellules). Bien que ces détails ne soient pas « réalistes » (il est douteux que des entités microscopiques reflètent la lumière comme des bulles de savon), ils augmentent bien la picturalité de l'image, en la rendant plus saturée (en ajoutant des propriétés syntaxiquement pertinentes). Cet exemple montre que la notion goodmanienne de saturation parvient à saisir quelque chose de la notion intuitive de « picturalité » : ajouter des reflets lumineux sur les cellules donne une impression de relief et de « réalité » qu'un schéma peu saturé n'a pas (quand bien même ce détail est tout à fait irréaliste).

³⁰ Il en résulterait bien un autre ensemble de marques, mais cet ensemble de marques serait une instance du même caractère : de même que différentes marques peuvent être des instances du mot « oculiste », défini comme la classe de ses instances, différentes marques peuvent être des instances du même schéma. C'est là une des différences entre les représentations picturales (saturées) et les représentations schématiques : en raison de leur saturation, les contraintes pour que deux ensembles de marques soient des instances de la même représentation picturale sont plus fortes. Il est difficile, par exemple, de trouver deux ensembles de marques *visiblement* différents qui soient les instances

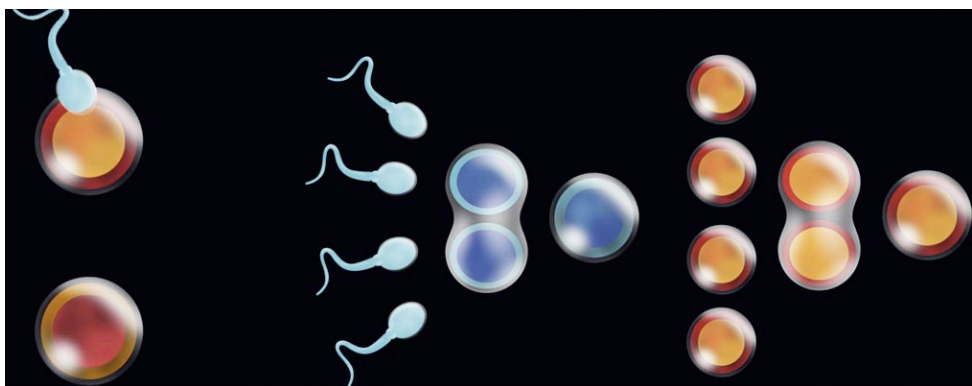


FIG. 25 – Schéma « pictural » de cellules reproductrices tiré d'un site de vulgarisation : les reflets lumineux sur les cellules ajoutent une dimension picturale à une représentation très schématique (très peu saturée) en augmentant sa saturation.

Prenons un autre exemple, qui n'est pas à strictement parler scientifique, mais qui va me permettre de souligner une propriété des représentations spatiales, et plus précisément de celles que j'appelle « schématiques », qui en fait des outils épistémiques particulièrement précieux. La figure 26 présente quatre représentations de la France. L'image satellite (en haut à gauche) est une représentation picturale. Les trois autres cartes peuvent être projetées sur cette image ; elles appartiennent bien (comme toutes les cartes géographiques) à la catégorie des schémas. La carte du relief de la France (en haut à droite) est nettement moins picturale (plus schématique) que l'image satellite : les pays frontaliers et la mer, entre autres, sont gommés. Des couleurs y ont été ajoutées qui indiquent l'altitude relative des différentes zones. Cette carte est cependant, si on la compare aux deux cartes de température juste en dessous, assez picturale (peu schématique) : elle est dense et très saturée. Les couleurs y sont réparties avec une grande précision, et la moindre modification entraînerait un changement dans le contenu de la carte ; c'est ce qui donne, d'ailleurs, l'impression de « voir » le relief comme sur une photographie. La carte des températures qui se trouve en bas à gauche est nettement plus schématique. Huit couleurs différentes, correspondant à des fourchettes de température, y sont réparties de telle sorte que l'on peut identifier rapidement les températures relatives des différentes zones. À la différence de ce qui est le cas dans la carte du relief, cette carte est peu saturée : seules les couleurs et leur emplacement relatif sont des propriétés syntaxiquement pertinentes. Enfin, la carte qui se trouve en bas à droite est aussi une carte des températures, sur laquelle ont été ajoutés des chiffres représentant les valeurs numériques des températures. Elle est un peu moins schématique que celle de gauche, dans la mesure où les pays frontaliers

de la même représentation picturale.

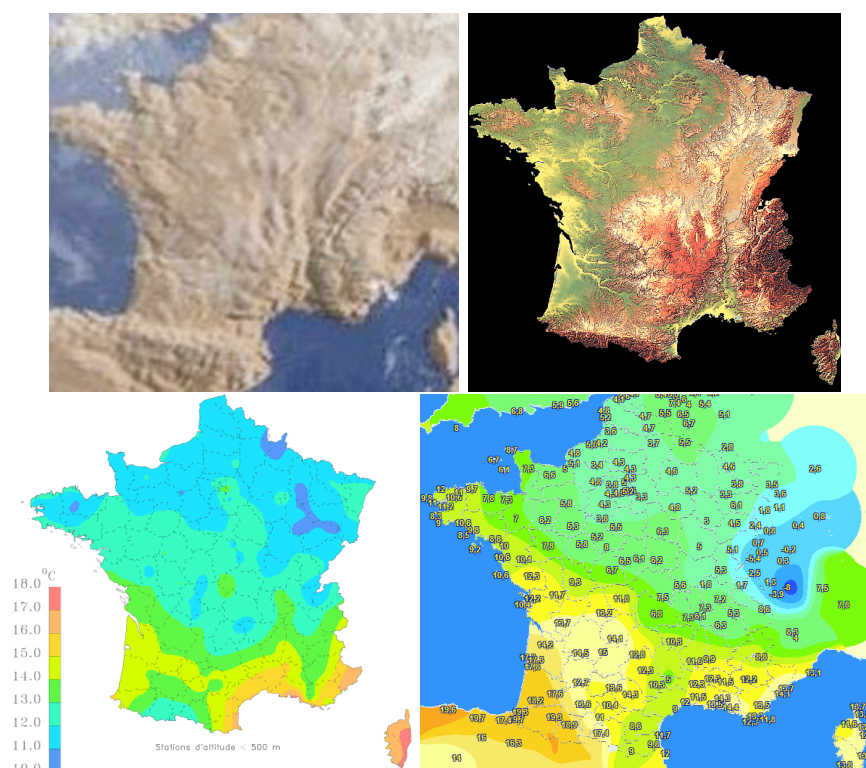


FIG. 26 – Cartes de France : température (en bas à gauche et en bas à droite), image satellite (en haut à gauche), relief (en haut à droite). Tirées du site de Météo France : <http://www.meteo.fr/> et de <http://www.populationdata.net> (pour la carte du relief).

sont représentés (la France n'y a pas été « découpée ») ; en revanche, des éléments linguistiques (des chiffres) y sont ajoutés, qui augmentent le caractère schématique de la carte, en y introduisant des éléments qui ne sont pas denses.

Les deux cartes du bas mettent en évidence une des propriétés remarquables des représentations spatiales soulignée par John Kulvicki (2009) : ce type de représentations permet de présenter *simultanément* des informations à différents niveaux d'abstraction, en attribuant à différentes propriétés syntaxiques des référents de diverses sortes. Les couleurs représentant les fourchettes de température me permettent de connaître la température approximative de chaque zone, ainsi que sa valeur relative dans les différentes zones : la relation entre les différentes teintes représente celle des températures. L'ajout des chiffres *sur la même carte* permet en outre d'en connaître les valeurs précises. Au moyen du langage, cela est impossible : une liste des valeurs précises des températures ne me donne pas *en même temps* accès aux valeurs relatives. Il me faut tirer des inférences et produire un nouvel énoncé portant sur ces relations ; ce nouvel énoncé, à son tour, ne contient plus l'information concernant

les valeurs précises de chaque point. Cette propriété des représentations spatiales (qui leur permet de présenter simultanément des informations à différents niveaux d'abstraction, ou des informations de types différents) leur confère manifestement une utilité authentiquement épistémique. Elle permet, par exemple, de mettre en relation différentes propriétés de la cible d'une représentation de manière beaucoup plus immédiate que ne le permettrait le langage. Pour prendre un exemple simple, la projection de la carte du relief (en haut à droite) sur la carte des températures (en bas à gauche) permettrait une mise en relation immédiate du relief et du climat de la France. Dans le cadre de la pratique scientifique, cette caractéristique des représentations spatiales, qui leur permet de présenter des hypothèses de façon immédiatement intelligible à propos de la relation entre différentes propriétés de la cible, est particulièrement précieuse (aussi bien pour l'enseignement que pour la recherche).

Comme les différents exemples que je viens de présenter le montrent, les schémas sont des représentations souvent composites, au sens où elles peuvent faire intervenir des éléments picturaux, ainsi que des éléments linguistiques. Une procédure courante consiste également à faire usage d'échelles différentes, en particulier dans l'enseignement et la vulgarisation. Cette procédure est particulièrement utile quand on cherche à présenter la structure fine de la cible (par exemple, celle de la cellule) tout en montrant la cible entière (la cellule). Les schémas de la figure 27, issus de sites de vulgarisation sur Internet, présentent ainsi respectivement une cellule entière et la double hélice de l'ADN et « entrent » progressivement dans leur structure fine, s'éloignant de plus en plus du pictural. Dans le schéma de droite, les bases de l'ADN sont représentées par des diagrammes moléculaires qui, comme je l'ai indiqué au début de cette section, ne sont ni denses ni saturés, et n'appartiennent pas à la catégorie des représentations picturales et schématiques (car les relations spatiales n'y tiennent pas lieu de relations spatiales, comme on le verra dans la prochaine section).

Soulignons pour finir que les représentations schématiques peuvent s'éloigner du pictural au point de ne plus du tout respecter les distances relatives comme, par exemple, dans la représentation de la France au moyen d'un hexagone (figure 28) ou d'un chromosome au moyen d'un bâton.³¹ De telles représentations sont cependant schématiques au sens où les relations topologiques de la cible sont reproduites au sein du schéma. La classe des représentations schématiques est donc très large. Elle inclut les schémas dont la projection sur une image picturale se fait presque sans distorsion

³¹ Notons d'ailleurs que, dans la figure 26, les deux cartes qui excluent les pays frontaliers distordent également certaines relations métriques : la Corse y est « rapprochée » du continent, ce qui explique que les deux cartes concernées (en haut à droite et en bas à gauche) soient plus étroites, à hauteur égale, que les deux autres.

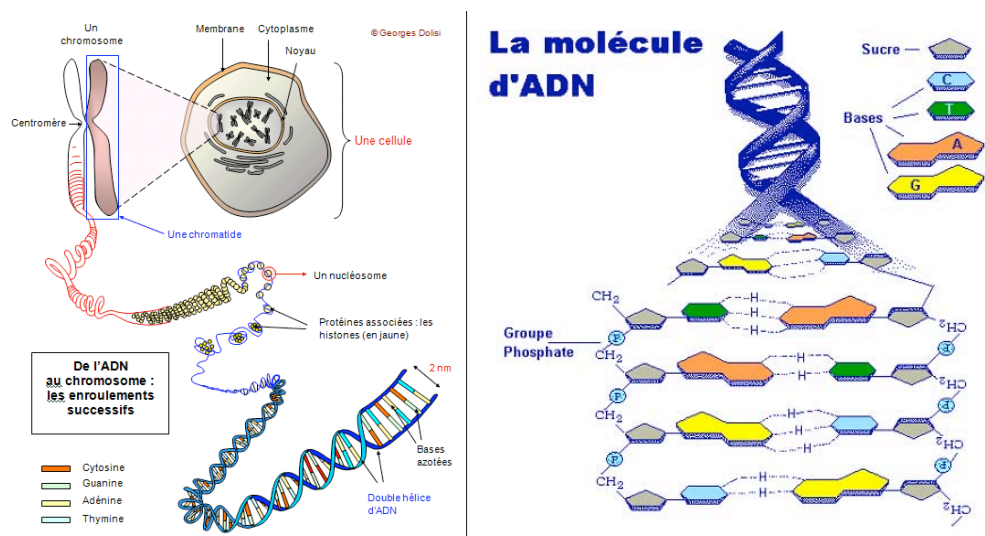


FIG. 27 — Schémas tirés de sites de vulgarisation. À gauche : du chromosome à l'ADN ([http ://www.medicopedia.net](http://www.medicopedia.net)). À droite : de l'ADN aux bases ([http ://www.futura-sciences.com/fr](http://www.futura-sciences.com/fr)).

(comme celui de la figure 24), les *cartes*³² représentant des zones géographiques, cartes qui peuvent elles-mêmes être plus ou moins saturées et donc plus ou moins picturales (de l'image satellite à l'hexagone représentant la France), ainsi que, par exemple, certains *tableaux* à double entrée dont les colonnes et les lignes correspondent aux longitudes et latitudes d'une zone géographique (comme ceux de la figure 32, que j'étudierai dans la section 2).

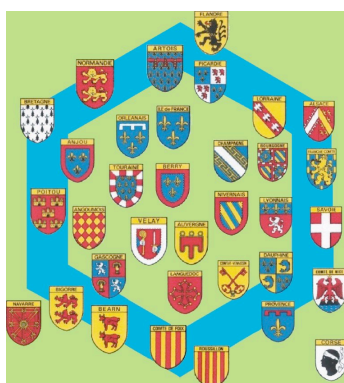


FIG. 28 — Hexagone représentant la France.

³²Voir (Casati et Varzi, 1999) pour une sémantique des cartes.

Avant de passer à l'examen des représentations diagrammatiques, récapitulons certains points importants de la définition des représentations schématiques, et de leur rapport aux représentations picturales.³³ En premier lieu, les représentations schématiques ont en commun avec les représentations picturales de respecter certaines relations topologiques de leurs cibles : les relations spatiales au sein du schéma représentent des relations spatiales au sein de sa cible, de telle sorte qu'un schéma doit toujours pouvoir être projeté sur une image picturale.

Cela a pour conséquence le fait que la différence entre une représentation schématique et une représentation picturale est affaire de degrés. En outre, on peut décrire le passage du pictural au schématique comme un processus de théorisation : le schéma exprime des hypothèses théoriques qui ne sont pas visibles dans l'image picturale, mais qui fournissent aux agents des clés de lecture des images picturales.

Ce processus de théorisation ou, si l'on veut, d'abstraction, se repère au caractère relativement peu saturé des schémas. D'une part, ils ont un moins grand nombre de propriétés syntaxiquement pertinentes. D'autre part, certaines de leurs propriétés syntaxiquement pertinentes ne se trouvent pas dans l'image picturale (par exemple, des couleurs et des symboles linguistiques y sont ajoutés, des traits y sont grossis).

1.3.4 Les représentations diagrammatiques

La catégorie de ce que j'appelle « diagrammes » est, comme celle des schémas, assez hétérogène. La propriété qui distingue en premier lieu les représentations diagrammatiques des représentations picturales et schématiques est que les relations spatiales en leur sein *ne représentent pas des relations spatiales*. Elles peuvent représenter des relations causales, temporelles, hiérarchiques, etc. Typiquement, une frise historique représente des relations temporelles au moyen de relations spatiales, et un organigramme représente des relations hiérarchiques. En ce sens, les diagrammes sont bien des représentations spatiales, dans la mesure où les relations spatiales y sont sémantiquement pertinentes – elles dénotent une propriété de leur référent, mais elles ne sont ni picturales ni schématiques, car les relations topologiques de leur référent n'y sont pas reproduites. À la différence des représentations picturales et schématiques, les diagrammes représentent souvent des relations abstraites et non des objets concrets : le mécanisme d'une réaction chimique, la synthèse d'une protéine, l'interaction de deux particules atomiques, l'évolution relative de deux variables, etc.

Au sein de la catégorie des diagrammes, les représentations peuvent être entièrement ou partiellement denses, et plus ou moins saturées. Les *graphes*, par exemple,

³³Cette définition constituera un élément important de mon étude de cas sur la cartographie génétique, au chapitre 9.

sont des diagrammes denses et très peu saturés. Les tableaux présentant spatialement des données numériques³⁴ ne sont ni denses ni saturés : ce qui en fait toutefois des diagrammes et les distingue de représentations linguistiques comme des listes de données numériques est leur organisation spatiale en lignes et en colonnes, qui a par elle-même une signification.

Les diagrammes, comme les schémas, sont souvent des représentations mixtes, faisant intervenir des éléments picturaux, ainsi que des éléments linguistiques. Ils servent à l'expression d'hypothèses théoriques, dont l'énoncé linguistique serait malaisé. Ainsi par exemple le « dogme central » de la biologie moléculaire est-il communément exprimé par un diagramme. Dans la figure 29, le diagramme de gauche, proposé par Crick (1970), fait intervenir très peu d'éléments picturaux. Il n'est ni dense ni saturé. Cependant, le sens des flèches, et l'agencement spatial des symboles en font bien une représentation spatiale, et non linguistique. La figure de droite, qui exprime aussi le dogme central de la biologie moléculaire, y introduit de nombreux éléments picturaux, ou tout au moins schématiques (comme la forme de la double hélice de l'ADN).

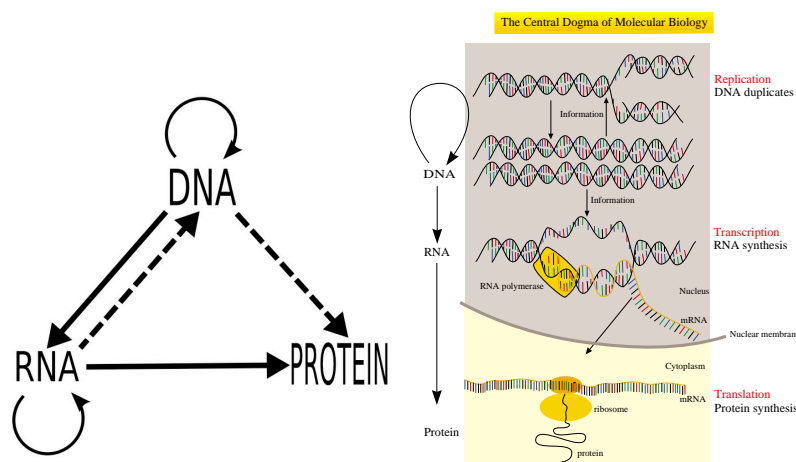


FIG. 29 – Dogme central de la biologie moléculaire. À gauche : diagramme de Francis Crick (1970). À droite, diagramme tiré du site personnel d'Andy Vierstraete, Université de Gand (<http://users.ugent.be/~avierstr/principles/centraldogma.html>).

Ces exemples suggèrent que la frontière entre schémas et diagrammes n'est pas toujours très nette : le diagramme de droite fait intervenir des éléments schématiques, de même que les schémas de la figure 27 emploient des éléments diagrammatiques (notamment les symboles représentant les bases de la molécule d'ADN). Cependant,

³⁴ À l'exception de ceux, mentionnés précédemment, au sein desquels les lignes et les colonnes correspondent aux latitudes et longitudes de la zone représentée, auquel cas il s'agit de schémas.

si je classe la représentation de la figure 29 dans la catégorie des diagrammes, c'est parce que les relations spatiales qui dominant représentent des relations causales (un mécanisme) et non pas des relations spatiales. Les éléments schématiques y sont secondaires, et y ont une valeur simplement iconique, au sens où ils permettent d'identifier les molécules d'ADN ; cependant, comme le diagramme de gauche le montre, on peut remplacer ces schémas par des symboles linguistiques sans modifier l'essentiel du contenu de la représentation. Ce n'est pas le cas dans la figure 27 : ce que ces schémas mettent en évidence, c'est bien la structure spatiale, à différentes échelles, de la cellule.

L'examen des représentations diagrammatiques appelle plusieurs remarques sur la notion de visualisation. On dit souvent que ces représentations permettent de visualiser des relations abstraites : un graphe permet ainsi de visualiser l'évolution relative de deux valeurs ; les diagrammes de la figure 29 permettent de visualiser le mécanisme de transcription de l'ADN. Cet emploi du terme « visualisation » n'a pourtant pas le même sens que quand on parle de visualisation dans le cas des représentations picturales et schématiques. Bien que les images picturales et les schémas présentent souvent des propriétés qui ne sont pas par elles-mêmes visibles, elles représentent, comme je l'ai dit, l'agencement spatial de ces propriétés par des relations spatiales ; en ce sens, elles permettent bien de visualiser la répartition spatiale d'une certaine propriété (par exemple, la température). En revanche, dans le cas des diagrammes, la notion de visualisation est *métaphorique*, dans la mesure où l'usage même de la notion d'*espace* est métaphorique.³⁵ C'est cet usage métaphorique de la notion d'espace, ainsi que des notions de *proximité* et de *distance*, qui permettent de parler de *cartes* dans le cas d'une représentation comme celle de la figure 30, où les relations de proximité et de distance qui y sont représentées ne sont pas spatiales. Cette « carte », obtenue à partir d'une enquête sur les recherches effectuées par les scientifiques sur Internet, représente leurs habitudes de recherche par discipline ; les disciplines y sont représentées comme des régions spatiales plus ou moins « proches ».

En outre, dans le cas des diagrammes, on emploie souvent le terme « visualisation » en deux sens différents, qui révèlent une particularité importante de ce type de représentation. Prenons l'exemple simple des graphes de la figure 18. Ils permettent de visualiser des relations abstraites (l'évolution temporelle relative des températures, de l'ensoleillement et des précipitations). Cependant, on peut aussi dire qu'ils per-

³⁵Pour une étude psychologique sur le rôle des métaphores spatiales dans l'utilisation de graphes et de diagrammes, voir notamment (Tversky, 2001, 2004). La dimension métaphorique des représentations spatiales est un des aspects du rôle des métaphores spatiales dans la vie cognitive, que le programme de recherche en linguistique cognitive de Georg Lakoff et Mark Johnson (voir par exemple Lakoff et Johnson, 1980) vise à étudier (comme je l'ai indiqué au chapitre 3, note 3, page 96).

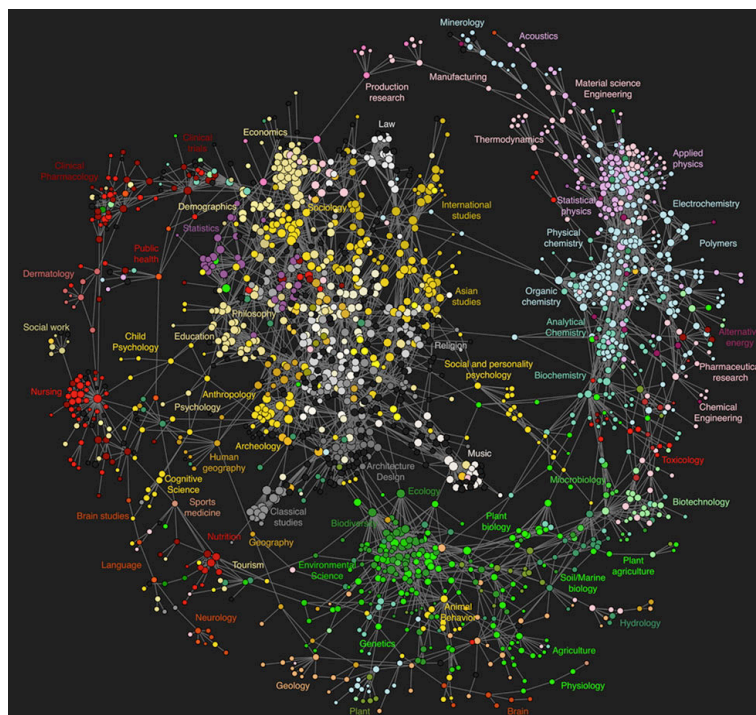


FIG. 30 — Carte des habitudes de recherche des scientifiques sur Internet ([http ://www.outilsfroids.net/news/une-cartographie-des-habitudes-de-recherche-des-scientifiques-sur-internet](http://www.outilsfroids.net/news/une-cartographie-des-habitudes-de-recherche-des-scientifiques-sur-internet)).

mettent de visualiser les données numériques elles-mêmes à partir desquelles on les a tracés, et dont ils sont une présentation spatiale. De façon plus générale, un graphe est toujours à la fois la *présentation spatiale du contenu d'une représentation linguistique* (liste de données, fonction mathématique, etc.) et la *représentation spatiale d'une relation non spatiale*.³⁶

Cette remarque peut sans doute être appliquée à la plupart des diagrammes. Ainsi, par exemple, le diagramme du dogme central de la biologie moléculaire représente un mécanisme au moyen de représentations spatiales ; il présente aussi, sous forme spatiale, le contenu de l'énoncé de ce « dogme » (de cette hypothèse théorique). Certes, comme je l'ai souligné à plusieurs reprises, *toutes* les représentations scientifiques expriment des hypothèses théoriques. Cependant, que ce je cherche à suggérer ici est la chose suivante : sans doute en raison du fait qu'ils représentent des relations

³⁶C'est ce que met en évidence le graphe présenté par Mach pour donner une représentation « à l'œil » de la loi du repos de Maupertuis (figure 31) : il suggère en effet que le graphe de l'équation de cette loi pourrait être généré en faisant glisser horizontalement une feuille blanche au niveau du poids $Q/2$: le résultat de son mouvement par rapport à la feuille serait le graphe lui-même.

abstraites et non spatiales, les diagrammes, et tout particulièrement les graphes, ont en général des contenus informationnels du même type que certaines représentations linguistiques. Contrairement à ce qui est le cas pour les schémas et les images picturales, il est relativement difficile d'envisager des informations qui pourraient être exprimées au moyen d'un diagramme mais qui seraient inexprimables au moyen du langage.³⁷ Alors qu'il semble difficile de dire qu'un schéma permet de « visualiser » les hypothèses théoriques qu'il exprime – et qui en sous-tendent, comme on l'a vu, la lecture –, il y a bien un sens dans lequel on peut dire qu'un graphe permet de visualiser la fonction ou les données numériques desquelles il est issu ; il contient les mêmes informations que la liste de ces données numériques (ou que l'expression linguistique de cette fonction). De même, le diagramme du dogme central de la biologie moléculaire ne contient pas d'autre hypothèse que l'énoncé linguistique de ce dogme. En revanche, comme la prochaine section va le montrer, ce type de représentation permet souvent, à contenu informationnel équivalent, de faciliter considérablement les inférences que les agents peuvent tirer par leur moyen.³⁸

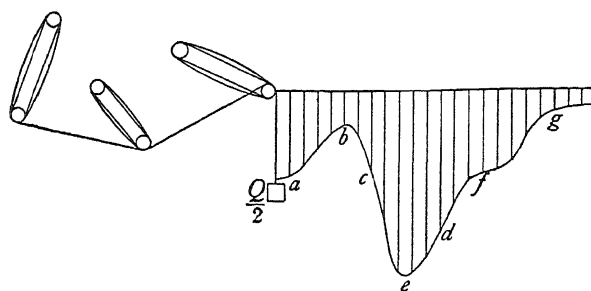


FIG. 31 – Diagramme présentant la loi du repos (Mach, 1883, p. 81).

Pour conclure cette ébauche de typologie, soulignons une fois de plus ce que j'ai signalé page 369, et qui est apparu à l'examen de ces quelques exemples : les types que je viens de dégager ne se trouvent que très rarement à l'état « pur ». D'abord, les images picturales elles-mêmes sont le fruit d'un travail impliquant lui-même de nombreuses hypothèses théoriques. Si, comme le dit Lynch (1988, p. 208), la trace de la « main » de l'artiste y est moins évidente, elles sont loin d'être des données brutes (elles ne sont pas de pures empreintes, analogues à la trace d'un pied sur le sol) et sont

³⁷C'est peut-être une des raisons pour lesquelles Goodman ne définit pas de critère permettant d'énoncer une distinction nette entre les langages et les représentations diagrammatiques dont la syntaxe est articulée.

³⁸Pour une analyse des types de raisonnements associés à l'usage des graphes et des diagrammes, voir (Tufte, 1983, 1990).

le résultat de nombreuses manipulations. Elles sont donc déjà, en ce sens, quelque peu schématiques. En outre, elles fonctionnent souvent, comme le montre l'exemple que j'ai emprunté à Lynch (figure 24) en interaction avec des représentations schématiques. Enfin, comme on l'a vu, les représentations schématiques et diagrammatiques sont la plupart du temps mixtes : les schémas font intervenir des symboles qui ne sont pas eux-mêmes schématiques (des symboles linguistiques ou diagrammatiques) ; les diagrammes font souvent intervenir des représentations schématiques, qui ont alors le rôle d'icônes.

Une des choses que j'ai cherché à mettre en évidence dans la présentation de ces exemples est que le classement de différents types de représentations sur la base des propriétés syntaxico-sémantiques des systèmes dans lesquels elles fonctionnent suggère que différents types de représentation sont adaptés à différents types de contenus. Les schémas et les images picturales représentent en général la structure spatiale d'entités concrètes ; les représentations diagrammatiques représentent des relations et des processus non spatiaux. En outre, dans chacune de ces deux grandes catégories (schématico-picturale et diagrammatique), des représentations de densité et de saturation différentes semblent adaptées à différents buts épistémiques. La possibilité de faire varier la saturation et la densité des représentations schématiques en fait des outils qui permettent, comme l'exemple des cartes de France l'a montré, de présenter simultanément des informations de différents types et à différents niveaux d'abstraction, et de « visualiser » leurs relations, ce qui est impossible au moyen du langage. En ce sens, la structure même de certaines représentations schématiques en fait des outils théoriques puissants, qui peuvent contenir un type et une quantité d'information qu'une description, même très longue, ne suffirait pas à exprimer.

Les diagrammes, quant à eux, semblent présenter un intérêt épistémique différent de celui des schémas. Sans permettre l'expression de contenus essentiellement différents de ceux des représentations linguistiques, ils permettent aux agents de tirer de inférences de manière beaucoup plus aisée qu'ils ne le pourraient au moyen d'un énoncé. Autrement dit, si un diagramme vaut mieux que mille mots, c'est en un sens différent de celui qui s'applique aux schémas. La prochaine section est ainsi consacrée à l'examen de la manière dont différentes représentations, tout en contenant les mêmes informations, n'en facilitent pas également l'accès aux agents.

2 La notion de format de représentation

L'analyse qui précède a cherché à montrer que l'utilisation de représentations de types divers dans la pratique scientifique obéit à des impératifs authentiquement épistémiques : les règles des systèmes symboliques dans lesquels différents ensembles de

marques fonctionnent comme représentations les rendent plus ou moins bien adaptés à l'expression de certains contenus. En particulier, les systèmes symboliques picturaux et schématiques permettent de présenter au moyen d'un même ensemble de marques des informations qu'une suite d'énoncés ne peut pas toujours exprimer.

L'importance de l'utilisation de représentations de types divers dans l'activité théorique ne se limite pourtant pas à cela. L'analyse des représentations au moyen de la notion de système symbolique permet de rendre compte du fait que, selon le système dans lequel il est interprété, *le même ensemble de marques ne contient pas les mêmes informations*. Je voudrais à présent m'intéresser aux cas où *différents ensembles de marques* contiennent (partiellement ou totalement) *les mêmes informations*.

Par exemple, une liste de données numériques et la présentation des mêmes données sous la forme d'un tableau (comme dans la figure 32, page 389) sont deux ensembles de marques qui, fonctionnant dans deux systèmes symboliques différents, contiennent strictement les mêmes informations. Pourtant, la présentation sous forme de tableau, *sans modifier les informations* contenues dans la liste, en rend certaines plus *facilement accessibles*. Ici, l'expression selon laquelle une image vaut mieux que mille mots prend un sens nouveau³⁹ : le tableau ne contient rien que le langage ne puisse exprimer, mais il *facilite les inférences* nécessaires à l'obtention de certaines informations.

Dans cette section, je me propose d'analyser les changements de représentation comme celui qui caractérise le passage d'une liste de données numériques à la présentation de ces données sous la forme d'un tableau *du point de vue des conséquences que ces changements entraînent pour les raisonnements des agents*. Le phénomène auquel je m'intéresse ici n'est donc pas tant la relation entre la structure d'une représentation et son contenu (étudiée dans la section précédente) que la conséquence de cette relation pour les raisonnements que des agents aux capacités cognitives limitées peuvent effectuer au moyen de cette représentation.

Je propose d'appeler les changements qui, sans (forcément) modifier le contenu informationnel d'une représentation, ont des conséquences sur les raisonnements des agents des *changements de format*. Ainsi, deux ensembles de marques qui, lues dans deux systèmes symboliques différents, ont le même contenu mais ne le rendent pas également accessible sont des représentations dans deux formats différents. Une liste et un tableau, l'expression mathématique d'une fonction et son graphe, l'énoncé du dogme central de la biologie moléculaire et le diagramme correspondant (figure 29) sont des exemples de représentations (au moins partiellement) équivalentes dans des formats

³⁹C'est le sens qu'elle a dans l'article de Simon et Larkin (1987) intitulé, « Pourquoi un diagramme vaut (parfois) mieux que dix mille mots » [*Why a diagram is (sometimes) worth ten thousand words*].

différents. Des représentations de même type (linguistique, diagrammatique, pictural) peuvent aussi, dans certains cas, être dans des formats différents. Par exemple, un graphe en coordonnées cartésiennes et un graphe en coordonnées polaires ne présentent pas les informations qu'ils contiennent de la même manière. Des équations exprimées dans des systèmes de coordonnées différents peuvent aussi, comme je le montrerai dans ce qui suit, être considérées comme des représentations dans des formats différents, dans la mesure où elles n'exigent pas de la part des agents les mêmes processus inférentiels.

En ce sens, les différences de format sont plus « fines » que les différences entre les types de représentations présentés dans la section précédente. Pour autant, il semble possible, de prime abord, de les caractériser au moyen des règles syntaxico-sémantiques des systèmes dans lesquels ces représentations doivent être interprétées. En effet, c'est parce qu'ils sont lus dans des systèmes différents que des ensembles de marques différents expriment le même contenu. Autrement dit, à tout changement de format semble bien correspondre un changement de système symbolique : deux graphes dans des systèmes de coordonnées différents ne sont pas interprétés en vertu des mêmes règles syntaxico-sémantiques.

Cependant, la notion de format n'est pas synonyme de celle de système symbolique : alors que la seconde rend compte, comme je l'ai dit, de la relation entre la structure d'une représentation et son contenu, la première désigne la relation entre la présentation de ce contenu et les opérations que les agents doivent effectuer pour accéder à ce contenu. Une analyse de la notion de format doit donc prendre en compte non seulement le système symbolique dans lequel une représentation est censée être interprétée mais également les opérations que des agents aux capacités cognitives limitées doivent effectuer pour accéder à son contenu. Autrement dit, un ensemble de marques contenant un certain nombre d'informations en vertu d'un certain système symbolique n'est pas dans le même format pour Dieu (dont on suppose les capacités cognitives illimitées), pour un ordinateur, et pour un humain, parce ces différents « agents » ne traitent pas les informations contenues dans cette représentation au moyen des mêmes opérations.

Dans ce qui suit, je propose dans un premier temps d'approfondir l'analyse de la notion de format et de montrer son importance pour l'activité théorique au moyen de plusieurs exemples (section 2.1). Ensuite, je cherche à en donner une définition précise en distinguant le contenu cognitivement accessible d'une représentation de son contenu informationnel (section 2.2).

2.1 Quelques exemples de l'importance du format des représentations

Afin de clarifier les intuitions qui sous-tendent la notion de format, prenons pour commencer un exemple simple, emprunté à John Kulvicki (2009). Les résultats d'un relevé de température sur une certaine zone géographique peuvent être présentés de différentes manières, comme le montre la figure 32. À gauche, ils sont présentés sous la

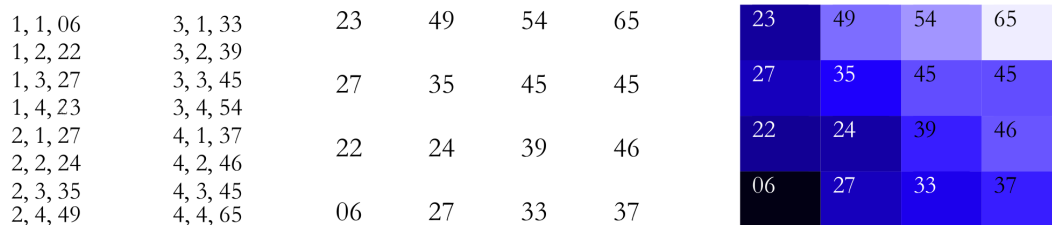


FIG. 32 – Résultats d'un relevé de température. Image tirée de (Kulvicki, 2009).

forme d'une liste de triplets de chiffres, les deux premiers représentant les coordonnées des différents points de la zone où les mesures ont été prises, et le troisième la valeur de la température correspondante. Les mêmes données peuvent être présentées sous la forme d'un tableau, où les emplacements des chiffres représentant les valeurs de la température respectent la distance relative des points où les mesures ont été prises : les tableaux du centre et de droite présentent spatialement les données de la liste.⁴⁰ Dans le tableau de droite, des couleurs correspondant à différentes fourchettes de température ont été ajoutées.

Les données présentées dans ces trois figures sont strictement identiques. Cependant, certaines informations sont plus facilement accessibles au moyen des tableaux. En préservant les relations topologiques entre les points de la zone représentée, ces derniers rendent en effet la comparaison entre les valeurs de la température en deux endroits différents beaucoup plus facile que ne le serait cette opération au moyen de la liste. L'ajout de couleurs dans le tableau de droite augmente encore la facilité d'obtention de certaines informations : l'intensité relative des tons bleus rend les valeurs relatives de la température immédiatement perceptibles.⁴¹ Un simple coup d'œil suffit

⁴⁰Selon les définitions données dans la section précédente, ces tableaux entrent dans la catégorie des schémas : par leur organisation en lignes en colonnes, qui correspondent aux coordonnées spatiales de la zone géographique représentée, ils représentent (très) schématiquement la répartition spatiale des températures sur cette zone en respectant les relations topologiques.

⁴¹Mon usage de l'expression « immédiatement perceptible » ne présuppose aucune hypothèse sur la nature de la perception. Je veux simplement dire que l'information correspondante n'est le fruit

à conclure que la température est plus élevée au nord-est de la zone qu'au sud-ouest, sans qu'il soit nécessaire de comparer les valeurs numériques des coordonnées. Obtenir cette information au moyen du tableau en noir et blanc exigerait un processus inférentiel un peu plus long, quoique pas très difficile pour un agent aux capacités cognitives normales, et reposerait sur des opérations différentes : la comparaison de valeurs numériques et non pas celle de tons de couleurs. En revanche, obtenir cette information au moyen de la seule liste serait autrement plus laborieux : il faudrait mémoriser et comparer entre elles les valeurs présentées sur différentes lignes. Une manière naturelle de procéder consisterait d'ailleurs à tracer un tableau, c'est-à-dire à *transformer* la liste en une autre représentation contenant les mêmes informations mais les présentant d'une manière plus adaptée à la tâche à accomplir.

Ce premier exemple met en évidence un phénomène bien connu et étudié dans les domaines des sciences cognitives et de l'intelligence artificielle (voir Simon et Larkin, 1987; Larkin, 1989; Kleinmuntz et Schkade, 1993; Zhang, 1997) : selon la manière dont les données sont présentées, les capacités d'un agent à résoudre un problème peuvent être réduites ou diminuées. Pour reprendre les termes de Simon et Larkin (1987), deux représentations peuvent être *équivalentes d'un point de vue informationnel* mais *différentes d'un point de vue computationnel* : tout en contenant les mêmes informations (puisque chacune peut être obtenue à partir de l'autre), elles ne rendent pas les différentes informations qu'elles contiennent également accessibles aux agents (elles ne facilitent pas les mêmes inférences).

Les formats de la liste et du tableau se différencient donc par le type de manipulation et d'inférence qu'ils exigent de la part des agents. Alors que, au moyen de la liste, il faut effectuer plusieurs inférences pour obtenir des informations concernant les valeurs relatives des températures, le tableau, *en vertu de son format*, nous en dispense, en effectuant pour ainsi dire « lui-même » ces inférences : il suffit d'y entrer certaines données suivant une procédure définie (disposer les valeurs correspondant aux températures de manière ordonnée, et, éventuellement, colorier les cases du tableau suivant certaines instructions) pour que, aussitôt, apparaissent les valeurs relatives des températures des différentes zones. Il n'est pas nécessaire d'y entrer les informations concernant les coordonnées des zones concernées ; elles sont exprimées par la forme même (la disposition spatiale) du tableau. Pour employer une métaphore informatique, le tableau est un ordinateur plus puissant que la liste ; par sa seule forme, il supplée aux raisonnements des agents.

Un des buts de ce chapitre est de montrer que ce phénomène a une grande impor-

d'aucune inférence, de même que, en présence d'une flaque d'eau, l'énoncé « il y a une flaque d'eau » n'est le résultat d'aucune inférence, à la différence de « il a plu » ou même de « de l'eau est tombée sur le sol ».

tance dans le cas des représentations théoriques. Autrement dit, je voudrais montrer que les différences de format ont des conséquences sur l'activité théorique en tant qu'elle consiste à développer et à appliquer des hypothèses théoriques, et non pas seulement sur la manipulation de données. Pour cela, je vais m'appuyer sur trois exemples ; les deux premiers me sont fournis par le cas des équations de la mécanique et le troisième par celui des diagrammes de Feynman. Je les présente brièvement dans la suite de la présente section. Puis, dans la section 2.2, j'approfondirai la notion de format au moyen d'exemples simples, pour enfin revenir, dans la section 3, aux exemples des équations de la mécanique et des diagrammes de Feynman.

Prenons, dans un premier temps, l'exemple (exposé au chapitre 2) des différentes formes sous lesquelles les équations de la mécanique classique peuvent être exprimées. Je n'affirme pas que ces différences se réduisent à des différences de format ; on a vu, dans la première partie de ce travail, que la diversité des formulations de la mécanique peut être étudiée sous plusieurs angles différents, selon que l'on insiste, par exemple, sur les différences de représentation mathématiques ou sur les différences conceptuelles. Je me propose, ici, de considérer les différentes formes d'équations (newtoniennes, lagrangiennes, hamiltoniennes) comme des représentations autonomes permettant de résoudre des problèmes de mécanique. Autrement dit, je fais momentanément abstraction de leur insertion dans un ensemble déductif de principes (la théorie) et je les considère comme des « morceaux » de théorie permettant de représenter des systèmes physiques ; dans cette perspective, on peut décrire leurs différences comme des différences de format. Selon le problème qu'il s'agit de résoudre, l'usage de l'une ou l'autre forme d'équations peut en effet rendre beaucoup plus aisées ou difficiles les deux étapes de résolutions d'un problème de mécanique – l'écriture des équations et leur intégration.

Considérons d'abord les équations newtoniennes ($\mathbf{F} = m\mathbf{a}$). Afin de les utiliser pour décrire un système, il faut y entrer un certain nombre d'informations. Une fois que ces informations y figurent, la manipulation de l'équation suivant certaines règles de transformation (celles du calcul différentiel) permet d'obtenir des solutions. Le format des équations newtoniennes exige que l'on y entre des données relatives aux forces, puisque ces dernières y apparaissent explicitement. Les forces font donc partie des données dont on doit disposer afin de pouvoir utiliser cette forme d'équation. Or, on a vu (chapitre 2, page 49) que dans le cas de certains systèmes contraints, l'identification et la spécification de chacune des forces maintenant les contraintes sont pratiquement impossibles. Dans ces cas, le format lagrangien ($\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0$) est beaucoup plus adapté. Ces équations contiennent implicitement les contraintes, puisque le système de coordonnées dans lequel elles décrivent le problème les prend en compte. Le format lagrangien consiste à présenter sous une forme explicite les

informations concernant l'énergie du système, plutôt que celles concernant les forces. Le fait que les forces ne figurent pas dans les données du problème n'importe donc pas.

Le cas de la mécanique nous offre un deuxième exemple de changement de format permettant de faciliter la résolution d'un problème. Les équations lagrangiennes n'admettent pas, dans certains cas, de solution analytique. Comme on l'a vu au chapitre 2 (section 3.1), un changement de coordonnées, correspondant à la transformation mathématique appelée « transformation de Legendre », permet d'obtenir, à partir d'une équation insoluble du deuxième ordre, deux équations hamiltoniennes du premier ordre dont l'intégration est aisée ($\frac{\partial H}{\partial p_i} = \dot{q}_i$ et $\frac{\partial H}{\partial q_i} = -\dot{p}_i$). C'est alors la seconde étape de résolution d'un problème, celle de l'intégration des équations, qui est facilitée par un changement de format.

Mon troisième exemple est celui de l'utilisation d'un type de représentation non linguistique en physique théorique : celui des diagrammes de Feynman.⁴² Richard Feynman présente pour la première fois sa méthode diagrammatique en 1948 dans le cadre de l'électrodynamique quantique, c'est-à-dire la branche de la théorie quantique des champs qui décrit l'interaction électromagnétique de particules chargées (photons et électrons). Pour déterminer l'état d'un système après interaction, connaissant l'état initial du système, on calcule l'amplitude de probabilité de toutes les façons de passer de cet état initial à l'état final. Ces amplitudes de probabilité sont déterminées à partir du lagrangien des champs associés aux diverses particules. En pratique, il est impossible de calculer exactement les différentes amplitudes de probabilité et les physiciens utilisent la méthode des « perturbations » : ils n'effectuent qu'un calcul approché de ces amplitudes, en considérant les termes d'interaction du lagrangien comme de petites « perturbations » afin de simplifier les calculs. Les solutions d'interactions complexes sont approchées au moyen de solutions moins complexes auxquelles on ajoute des corrections d'ordre plus élevé.⁴³ La méthode de Feynman consiste à représenter par un diagramme chacune des contributions perturbatives d'un processus donné (c'est-à-dire les différentes façons de passer d'un état initial à un état final) à un ordre donné de perturbations. Feynman présente initialement⁴⁴ ses diagrammes

⁴²Pour les détails historiques et techniques, voir (Kaiser, 2000, 2005).

⁴³Les termes perturbatifs étant petits relativement aux autres termes, on considère qu'ils n'ont qu'une faible incidence sur les champs associés aux particules, qui peuvent alors être développés en une somme infinie de termes plus simples à étudier. En outre, la contribution de ces termes étant de plus en plus faible, on peut ne garder que les premiers termes de la série et négliger tous les autres.

⁴⁴Le but de Feynman consiste en outre à trouver une solution au problème des divergences causées par la « polarisation du vide » et les effets d'« auto-interaction » de l'électron qui produisent des quantités infinies dans les calculs (voir Kaiser, 2005, pp. 28-37), problème qu'il parvient à traiter au moyen d'une méthode de « renormalisation ». Cette procédure consiste à résorber les infinis dans

comme des « aide-mémoire » (Kaiser, 2000, p. 52) permettant d'effectuer des calculs perturbatifs d'ordre élevé sans confondre ou oublier les termes correspondant aux différents comportements possibles des particules virtuelles (voir figures 33 et 34). Dans la mesure où chacune des équations correspondant à une contribution possible occupe à elle seule plusieurs lignes, il est pratiquement impossible d'effectuer ces calculs au moyen des seules formules mathématiques.

certaines paramètres du lagrangien, comme la masse et la charge : on ajoute à la masse et à la charge, dans le lagrangien, des « contre-termes », c'est-à-dire des termes infinis qui compensent leurs infinis et qui sont ajustés pour donner la masse et la charge observées. Les contre-termes compensent les infinis introduits par les intégrations sur les impulsions des particules virtuelles, et conduisent à des amplitudes de probabilité finies. Voir (Kaiser, 2005, pp. 37-43). Feynman parviendra à trouver le moyen de couper les intégrales avant qu'elles divergent vers l'infini en 1949 (Kaiser, 2005, p. 43).

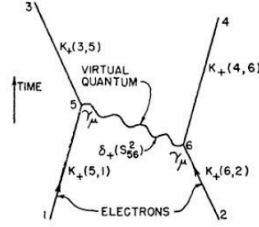


FIG. 33 – Ce diagramme représente l'interaction la plus simple entre deux électrons : un des électrons émet un photon virtuel, qui est absorbé par l'autre électron. Il sert ainsi à calculer la contribution de cette interaction au calcul complet. L'électron de gauche a une certaine probabilité de se mouvoir du point 1 au point 5, notée $K_+(5,1)$. L'autre électron a aussi une certaine probabilité de se mouvoir de 2 en 6, notée $K_+(6,2)$. Ce second électron peut alors émettre un photon virtuel en 6. Le photon a une certaine probabilité de se mouvoir de 6 en 5, notée $\delta_+(s_{56}^2)$. En arrivant au point 5, le premier électron peut absorber le photon. La probabilité qu'un électron émette ou absorbe un photon virtuel est notée $e\gamma_\mu$. Ayant perdu une partie de son énergie et de son impulsion, l'électron de droite se mouvrait alors de 6 en 4. Celui de gauche, ayant absorbé le photon et gagné de l'énergie et de l'impulsion, irait de 5 en 3. Ce diagramme correspond à l'équation : $K^{(1)}(3,4;1,2) = -ie^2 \int \int K_{+a}(3,5)K_{+b}(4,6)\gamma_{a\mu}\gamma_{b\mu} \times \delta_+(s_{56}^2)K_{+a}(5,1)K_{+b}(6,2)d\tau_5 d\tau_6$. (Feynman, 1949, p. 772)

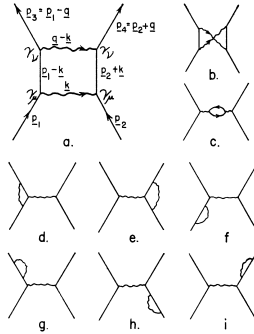


FIG. 34 – Différentes interactions simples entre deux électrons. Par exemple, dans la figure c, le premier électron émet un photon virtuel, qui se désintègre à mi-chemin en une paire électron-positron virtuelle. Cette paire se transforme à son tour en un nouveau photon qui continue sa route jusqu'au second électron. Pour chacun de ses diagrammes, Feynman écrit la contribution mathématique correspondante, en notant K_+ chaque branche du mouvement d'un électron, δ_+ chaque échange de photon, $e\gamma_\mu$ chaque sommet où les lignes d'un photon et d'un électron se croisent. (Feynman, 1949, p. 787)

Un an plus tard, Freeman Dyson (1949a,b) démontre l'équivalence des diagrammes de Feynman et des dérivations mathématiques proposées au même moment par Julian Schwinger (1948a,b) en guise de schème calculatoire utilisable en électrodynamique quantique. En outre, Dyson rend les méthodes de Feynman « publiquement utilisables »⁴⁵ en codifiant les règles permettant de construire les diagrammes sur la base d'une correspondance bi-univoque entre les propriétés du diagramme et les expressions mathématiques particulières.

Ainsi codifié, l'usage des diagrammes de Feynman – qui avaient paru mystérieux à la plupart des physiciens quand ce dernier les traçait « intuitivement » sans en expliciter les règles – devient rapidement une pratique courante en théorie quantique des champs. En outre, les diagrammes vont « survivre » à l'échec de la théorie des perturbations dans le cadre de laquelle la méthode diagrammatique semblait pourtant prendre tout son sens : progressivement, ils vont être utilisés et interprétés de plusieurs manières différentes et dépasser le cadre théorique dans lequel ils ont été inventés (voir Kaiser, 2005). Je reviendrai sur ce point dans la dernière section de ce chapitre. Pour l'heure, notons qu'ils ont joué un rôle central dans les recherches en physique des hautes énergies de l'après-guerre, et font partie aujourd'hui encore de l'apprentissage des étudiants en physique théorique, jouant le rôle de ce que Kuhn (1970a, 1969) appelle des « cas exemplaires » (voir ci-dessus, chapitre 6). Initialement introduits en guise d'outils mnémotechniques destinés à aider les physiciens à conduire des calculs difficiles, ils deviennent finalement d'authentiques *outils théoriques*, indépendants des hypothèses de l'électrodynamique quantique. Comme Kaiser (2005, p. 75) le suggère, Dyson est parvenu à démontrer « l'équivalence mathématique [...] entre les formalismes de Schwinger et de Feynman », mais « en aucun cas [...] leur équivalence conceptuelle ». Cela signifie que, malgré l'inter-déductibilité des formules mathématiques et des diagrammes, ces deux types de représentation ne jouent pas le même rôle dans les processus inférentiels des physiciens. Selon que l'on utilise les équations ou les diagrammes, on ne résout pas les problèmes de la même manière et, ce faisant, on n'a pas la même compréhension des hypothèses théoriques qu'ils expriment.

Dans tous les exemples que je viens de prendre, des représentations dont l'équivalence peut être prouvée mathématiquement s'avèrent avoir des conséquences différentes pour la résolution de problèmes, et même, dans certains cas, pour le développement des hypothèses théoriques qu'elles expriment. Selon la compréhension intuitive que l'on a de la notion de « format », suggérée par l'exemple de la figure 32, il semble sensé de dire que le cas de la mécanique comme celui des diagrammes de

⁴⁵Dyson, Lettre à ses parents, 4 décembre 1948, cité dans (Kaiser, 2005, p. 77).

Feynman nous fournissent des exemples de différences de format. Les équations newtoniennes, lagrangiennes et hamiltoniennes d'une part, les diagrammes de Feynman et les équations correspondantes d'autre part, sont dans des formats différents. Bien qu'elles contiennent, au moins partiellement, les mêmes informations (puisqu'elles sont toutes, au moins partiellement, inter-déductibles), elles ne les transmettent pas de la même manière ; les différentes informations qu'elles contiennent ne sont pas également accessibles à des agents cognitifs. Les construire et les utiliser ne consiste pas dans les mêmes opérations : identifier et spécifier les forces appliquées à un système en manipulant des vecteurs n'est pas le même processus qu'identifier le lagrangien d'un système. De même, dessiner un diagramme permettant de visualiser les différentes quantités à se rappeler au cours d'un calcul ne revient pas au même que faire ce calcul au moyen d'une formule mathématique. Ces différentes représentations ne jouent pas le même rôle dans les processus inférentiels des agents.⁴⁶

On retrouvera ces deux exemples dans la dernière section du présent chapitre. Auparavant, je souhaite examiner de plus près la notion de format elle-même, sur la base d'exemples simples. Une première définition intuitive de cette notion pourrait être la suivante : le format d'une représentation est la manière dont elle transmet l'information qu'elle contient à son utilisateur. Deux représentations dans des formats différents ne facilitent pas les mêmes inférences chez les agents. Un changement de format implique donc un changement dans le rôle inférentiel qu'une représentation joue pour ses utilisateurs. La prochaine section vise à rendre compte de cette « manière » de transmettre les informations, et de son influence sur les processus inférentiels des agents.

2.2 Obtenir des informations : coût cognitif et types d'inférences

Comme je l'ai dit dans l'introduction de la section 2 (page 388), une définition de la notion de format, à la différence de celle de la notion de système symbolique, doit rendre compte de la relation entre la présentation du contenu d'une représentation et les opérations que les agents doivent effectuer pour accéder à ce contenu. En conséquence, la définition du format d'une représentation particulière doit prendre en compte non seulement le système symbolique dans lequel elle est censée être lue mais également les opérations que ses utilisateurs peuvent et doivent effectuer pour accéder à ce contenu.

⁴⁶Les processus cognitifs ne sont pas réductibles à des processus inférentiels. Dans le cas du tableau en couleur, par exemple, certaines capacités de perception sont mises en œuvre qui n'entrent pas en jeu dans l'utilisation du tableau en noir et blanc. Cependant, ces différences d'opérations cognitives ont des conséquences sur les processus inférentiels nécessaires à traiter l'information. En conséquence, je parlerai indifféremment de processus cognitifs et de processus inférentiels.

Définir la notion de format implique en premier lieu d'affiner l'analyse du contenu d'une représentation et de distinguer entre son contenu informationnel (défini à la section 1.2 du présent chapitre) et ce que je propose d'appeler son « contenu cognitivement accessible ». Après avoir précisé cette distinction (section 2.2.1), je définirai (section 2.2.2) ce que j'entends par l'expression « accessibilité d'une information ». Cette analyse me conduira enfin à proposer une définition de la notion de format (section 2.2.3).

2.2.1 Contenu informationnel et contenu cognitivement accessible

Revenons un instant à la définition que j'ai donnée précédemment du contenu informationnel d'une représentation : c'est l'ensemble des informations qu'un agent maîtrisant le système symbolique dans lequel elle fonctionne – et en vertu duquel elle contient cette information – peut en principe en tirer. La définition du contenu informationnel d'une représentation dépend donc exclusivement de ses propriétés perceptibles et du système dans lequel elle fonctionne (voir ci-dessus, page 366).

Le fait que cette représentation contienne ces informations dépend, certes, de l'existence d'un agent et d'un contexte d'utilisation fixant le système symbolique adéquat ; cependant, la définition du contenu informationnel est relative à un *agent idéal*, maîtrisant parfaitement le système et dont les capacités cognitives sont illimitées. Cette définition ne prend en compte ni les opérations concrètes qu'un agent aux capacités limitées doit effectivement faire, ni l'ordre dans lequel les différentes informations peuvent être obtenues, ni enfin la relative difficulté des opérations nécessaires à l'obtention de ces différentes informations. Les différences de format n'ont pas de conséquence sur les informations qu'un agent aux capacités cognitives illimitées peut tirer d'une représentation, ni sur la manière dont il les obtient (puisque'il n'a aucun effort à faire) : les différentes représentations de la figure 32 ont strictement le même contenu informationnel.

En revanche, on a vu que, pour des agents aux capacités cognitives limitées, des représentations dans différents formats ne facilitent pas les mêmes inférences. Les différentes informations contenues dans une représentation ne sont pas également faciles à obtenir. Certaines sont parfois pratiquement inaccessibles à un humain normal sans l'aide d'instruments externes. Par exemple, sans l'aide d'un ordinateur, il m'est totalement impossible, à partir des lignes de code correspondant à la photographie digitale d'un chien, d'en tirer des informations concernant l'aspect de ce chien. Moins dramatiquement, si un humain normal peut tirer les mêmes informations des différentes représentations de la figure 32 (page 389) sans l'aide d'un ordinateur (mais peut-être dans certains cas d'un papier et d'un crayon), ce n'est pas au prix du même

effort. Certaines informations sont moins accessibles que d'autres au sein de chacune de ces représentations ; en outre, chacune de ces représentations ne rend pas les mêmes informations également accessibles que les deux autres.

Du contenu informationnel d'une représentation, je propose de distinguer son *contenu cognitivement accessible* qui dépend de la *relative accessibilité des différentes informations* constituant son contenu informationnel. De même que le contenu informationnel d'une représentation est déterminé par son système symbolique (et par les propriétés perceptibles de cette représentation), son contenu cognitivement accessible est déterminé par son format (et par ses propriétés perceptibles).

Pour clarifier ce que j'entends par là, prenons un exemple simpliste. Imaginons deux représentations, *A* et *B*. *A* contient les informations *x* et *y*. Je peux obtenir ces différentes informations sans l'aide d'un outil extérieur, mais je ne peux obtenir *y* qu'après avoir obtenu *x*. Inversement, *B* contient *x* et *y*, mais ne me permet d'obtenir *x* qu'après avoir obtenu *y*. Dans un tel cas, les contenus cognitivement accessibles de *A* et de *B* sont différents. *A* et *B* ne sont donc pas dans le même format.

Supposons, pour continuer avec cet exemple simpliste, qu'il existe des agents cognitifs différents de moi pour qui la situation soit inversée. Dans ce cas, *A* a pour moi le même contenu cognitivement accessible que *B* pour ces agents. Donc, *A* est dans le même format pour moi que *B* pour ces agents. En ce sens, la notion de format est relative aux raisonnements que les utilisateurs des représentations sont capables de conduire. Pour reprendre les termes de Simon et Larkin (1987, p. 67), quand il s'agit d'évaluer l'équivalence *computationnelle* de deux représentations, et non pas leur équivalence *informationnelle*, il faut prendre en compte à la fois la manière dont les données sont structurées [*data structures*] mais aussi la nature des processus qui en effectuent le traitement [*processes that operate upon them*].⁴⁷ Cette idée repose sur une analogie avec le sens que la notion de format a en informatique. Le format utilisé pour stocker des données sur un support informatique définit en effet les procédures de *codage* et de *traitement* des données. Dans la mesure où deux processeurs différents ne sont pas toujours susceptibles d'appliquer les mêmes algorithmes, c'est-à-dire de suivre les mêmes règles de transformations des données stockées, le même ensemble de données, *stocké sous la même forme*, peut être utilisable par un processeur et non par un autre. Selon le type d'information recherchée, et selon le système de traitement des données dont on dispose, tel ou tel codage s'avérera plus ou moins adapté.

Suivant cette analogie, la définition du format d'une représentation devrait pouvoir rendre compte de la manière dont les « processeurs » particuliers que sont les

⁴⁷Simon et Larkin (1987, p. 68) renvoient sur ce point à Anderson (1984), qui montre que la distinction entre différentes représentations ne doit pas être fondée sur les marques elles-mêmes, mais sur les opérations que l'on effectue sur elles.

agents accèdent aux différentes informations contenues dans cette représentation. Il convient donc de préciser, dans la prochaine section, ce que j'entends par « accessibilité d'une information ».

2.2.2 Accessibilité d'une information

Comme je l'ai suggéré dans ce qui précède, la relative accessibilité d'une information dans une représentation dépend *à la fois* du système symbolique en vertu duquel elle y est contenue et des capacités cognitives des agents. Afin de comprendre cela, commençons par préciser en quoi consiste l'accessibilité *immédiate* d'une information dans une représentation. Intuitivement, cette expression correspond à l'idée selon laquelle cette information figure de manière explicite dans la représentation ; autrement dit, l'obtention de cette information ne requiert aucun processus inférentiel. Le fait qu'une information soit ou non immédiatement accessible dans une représentation semble dépendre, en premier lieu, de certaines propriétés syntaxico-sémantiques de cette représentation. En définissant ces propriétés, on verra comment les capacités propres des agents entrent en jeu dans la détermination de l'accessibilité relative des différentes informations. J'emprunte la définition qui suit à John Kulvicki (2009).

Selon John Kulvicki, une information est immédiatement accessible si elle est (i) présentée sous une forme extractible, (ii) syntaxiquement saillante, et (iii) sémantiquement saillante. La notion d'extractibilité est destinée à rendre compte de l'idée selon laquelle aucune inférence, même très rapide, n'est nécessaire à l'obtention de cette information. Kulvicki en propose la définition suivante : une information figure sous une forme *extractible* s'il existe « une caractéristique non sémantique de la représentation en vertu de laquelle cette représentation contient cette information et aucune autre information plus spécifique ». Cela signifie qu'une des propriétés syntaxiquement pertinentes de la représentation est responsable de cette information, et de rien d'autre. Par exemple, les régions rouges d'un doppler indiquent la présence d'un temps orageux d'une certaine intensité, et rien d'autre. Le fait d'être rouge, à lui seul, ne dit rien de l'emplacement de cette perturbation ; c'est l'emplacement relatif de cette couleur qui est responsable de l'information concernant l'emplacement de la perturbation. Ainsi, dans la figure 32, l'information concernant les valeurs relatives des températures de la zone représentée figure de manière extractible dans les tableaux et non dans la liste : l'emplacement relatif des chiffres (ou des couleurs) est une propriété syntaxiquement pertinente des tableaux en vertu desquels ils nous indiquent l'emplacement relatif des températures dans la zone représentée.

La notion d'extractibilité ne dépend pas, à strictement parler, des capacités cognitives des agents. L'information est ou n'est pas présente sous forme extractible. En

revanche, les notions de saillance syntaxique et sémantique, comme on va le voir, sont relatives aux agents. Saillances syntaxique et sémantique correspondent, pour le dire en des termes métaphoriques, à la « visibilité » d'une information.

Une information extractible est *syntactiquement saillante* si la propriété responsable de cette information est perceptuellement saillante. Comme on va le voir, la saillance syntaxique d'une information dépend des propriétés perceptibles de la représentation et des capacités perceptives de l'agent. Par exemple, une carte représentant des températures au moyen de différents tons verts et rouges rend certaines informations syntaxiquement saillantes pour un agent aux capacités perceptives normales (mais pas pour un daltonien). Selon le contraste des couleurs, ces informations sont plus ou moins saillantes. Imaginons un graphe identique à celui de la figure 18, à ceci près que les trois courbes sont dans la même couleur. Le contenu informationnel est inchangé. Les mêmes informations sont extractibles. En revanche, certaines sont moins saillantes syntaxiquement : il faut fournir un effort plus grand pour ne pas se tromper dans la comparaison des points. Un étourdi pourrait ne pas parvenir à en tirer certaines informations (en revanche le graphe en couleur ne facilite pas les inférences d'un agent qui ne perçoit pas les couleurs). Pour prendre un exemple encore plus trivial, un texte écrit dans des caractères de très petite taille rend l'information qu'il contient moins saillante syntaxiquement⁴⁸ qu'un texte écrit en gros caractères.

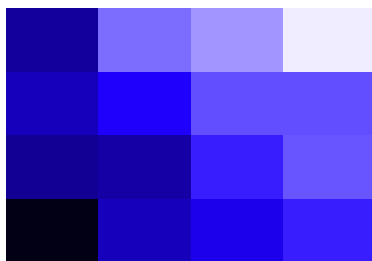


FIG. 35 – Image tirée de (Kulvicki, 2009)

Parfois, augmenter la saillance syntaxique d'une information s'accompagne d'une perte de données. Par exemple, dans le tableau de la figure 35, on peut considérer que l'absence des chiffres correspondant aux valeurs précises de la température de différentes zones (voir figure 32) évite de « brouiller » la lecture de l'image par des informations non pertinentes si ce que l'on cherche à connaître est la valeur relative

⁴⁸Il semble difficile de dire qu'un texte écrit présente des informations de manière extractible, en vertu de la définition qui vient d'être donnée de cette notion. Mais on peut dire que les propriétés syntaxiques permettant d'accéder à différentes informations sont moins saillantes syntaxiquement si le texte est écrit petit.

de la température des différentes zones, et non sa valeur exacte. L'absence des chiffres rend en ce sens les propriétés syntaxiques responsables de l'information recherchée (ces propriétés, ici, sont les différentes teintes de bleu) plus saillantes syntaxiquement.⁴⁹ Cela s'accompagne, cependant, de la perte des informations concernant les valeurs exactes de la température.

Pour prendre un autre exemple, imaginons que l'on projette la carte du relief de la France de la figure 26 sur la carte des températures indiquant les valeurs précises (celles qui se trouve en bas à droite). Le résultat serait tellement saturé qu'il deviendrait difficile de distinguer les propriétés syntaxiques responsables des différentes informations (relief et température) ; pour comparer le relief et la température de la France, on préférera projeter la carte du relief sur une carte ne comportant *que* les chiffres correspondant aux valeurs précises de la température, ou sur une carte ne comportant *que* les couleurs correspondant aux valeurs relatives. Dans les deux cas, la perte de certaines informations augmente la saillance syntaxique des autres.

La *saillance sémantique* d'une information dépend de la relation entre les propriétés perceptibles de la représentation et les données auxquelles elles correspondent : une carte de température où les tons rouges représentent des zones froides et les tons bleus des zones chaudes présente l'information relative à la température de manière moins saillante sémantiquement (mais pas syntaxiquement), étant donné nos conventions (et nos habitudes), que l'inverse.⁵⁰ De même, une frise historique où la flèche du temps irait de droite à gauche présenterait l'information relative à l'ordre des événements représentés de manière moins saillante sémantiquement. La saillance sémantique, comme la saillance syntaxique, est relative aux agents. Certains types de représentation, dont la lecture exige une certaine expertise de la part de l'agent, ne présentent pas l'information qu'elles contiennent de façon également saillante sémantiquement pour différents individus.

Saillances syntaxique et sémantique dépendent donc des capacités cognitives des utilisateurs et de leur maîtrise des règles syntaxico-sémantiques du système, et elles sont affaires de degré. Une information extractible mais peu saillante est moins accessible qu'une information extractible et saillante. L'effort à fournir pour l'obtenir est plus grand. Cet effort peut être un effort d'ordre perceptif, pour la saillance syntaxique : par exemple, une carte de température où les tons sont peu différenciés

⁴⁹L'augmentation de la saillance syntaxique par l'élimination de propriétés non pertinentes, qui brouillent la lecture, correspond à ce que Goodman appelle une baisse de la saturation de l'image (voir page 359).

⁵⁰Voir (Tversky, 2001, 2004) pour une étude psychologique sur l'universalité de certaines caractéristiques sémantiques des graphes et des diagrammes (elle montre notamment que les différentes significations attribuées à la verticalité se retrouvent dans la plupart des cultures).

demande un effort visuel de la part de l'utilisateur, plus susceptible de se tromper dans la lecture de la carte que quand le contraste est plus fort. En ce qui concerne la saillance sémantique, l'effort est d'une autre nature : si je suis habituée à lire la couleur rouge comme signifiant « chaud » et que je suis confrontée à une carte où les tons rouges représentent les zones froides, il me faut opérer une inférence de « rouge » à « froid » alors que, dans le cas d'une carte « normale » (qui représente les zones chaudes par des tons de rouge), on peut considérer que l'opération équivalente (de « rouge » à « chaud ») est immédiate, au sens où elle ne repose sur aucune inférence (du moins sur aucune inférence consciente et volontaire). On peut supposer que, avec l'habitude, l'effort inférentiel nécessaire à la lecture de la carte « inversée » soit réduit, au point de devenir quasiment nul.

La *relative accessibilité d'une information* pour un agent dépend, en premier lieu, de la saillance syntaxique et sémantique des propriétés perceptibles de la représentation responsables d'une certaine information. Dans le cas où cette information est celle que l'on cherche (autrement dit, dans le cas où l'information recherchée figure sous une forme extractible), son accessibilité dépend uniquement de la saillance syntaxique et sémantique des propriétés qui en sont responsables. Dans le cas où l'information recherchée ne figure pas sous une forme extractible, l'obtenir implique de mettre en œuvre un processus inférentiel à partir des informations disponibles (sous forme extractible). Ce processus exige de la part de l'agent la maîtrise d'un certain nombre de règles que j'ai présentées précédemment par analogie avec les algorithmes mis en œuvre par les ordinateurs pour traiter des données. De même que ces algorithmes peuvent être plus ou moins longs, les processus inférentiels permettant d'obtenir une information peuvent être plus ou moins difficiles. L'accessibilité d'une information dépend donc non seulement de la saillance syntaxique et sémantique de certaines informations, mais aussi de la *difficulté relative* de ce dernier processus.

Prenons de nouveau l'exemple des graphes de la figure 18 (page 362) : lire la valeur de la température au mois de juillet, de même qu'évaluer son évolution globale de février à juillet, sont des tâches aisées pour un agent moyen. Ces informations sont immédiatement accessibles. Donner la valeur exacte de la différence de température entre le mois de juillet et celui de décembre, quoique assez facile, implique d'identifier les valeurs et de faire la soustraction. Si ces valeurs étaient très précises, avec un grand nombre de décimales, cela pourrait exiger qu'on les écrive sur un papier afin de faire la soustraction. En revanche, donner la valeur exacte de la différence de température entre le mois de janvier et le mois de septembre exige plusieurs opérations. Dans le cas d'un graphe gradué, cela ne serait pas très difficile. Dans le cas du graphe de la figure 18, qui n'est pas gradué, cela implique de mesurer, au moyen des quelques informations précises disponibles, la valeur des températures relative à une unité de

distance, puis de mesurer la distance aux mois de janvier et de septembre, et enfin de faire la soustraction. Cela peut se faire au moyen de différentes méthodes. On peut par exemple ajouter des graduations au graphe, ce qui rend certaines informations accessibles. On peut aussi tracer une parallèle à l'axe des abscisses à la hauteur de la température au mois de janvier et mesurer la différence entre la hauteur de ce point et celui représentant la température en septembre. On pourra alors comparer par une règle de 3 la longueur de ce segment avec celui qui sépare les 3°C de décembre des 25°C de juillet. La proportion permettra de connaître la différence entre janvier et septembre. Toutes ces manipulations sont autorisées par le système symbolique dans lequel fonctionne le graphe, et on peut en inventer beaucoup d'autres.

Supposons pour l'instant que l'on puisse mesurer précisément les *efforts* nécessaires à l'identification, au sein d'une représentation, des propriétés pertinentes, ainsi que la *difficulté* des processus inférentiels permettant, à partir des informations ainsi obtenues, d'en obtenir d'autres, qui ne figurent pas sous forme extractible. L'ensemble définirait le *coût cognitif* nécessaire à l'obtention de ces informations. L'accessibilité relative d'une information est donc inversement proportionnelle au coût cognitif nécessaire à son obtention.

En vertu de ce que j'ai suggéré précédemment (page 398), la définition du format d'une représentation doit pouvoir rendre compte du coût cognitif nécessaire à l'obtention des différentes informations qu'elle contient. Ce n'est cependant pas tout. Il ne suffit pas de rendre compte de la difficulté du processus permettant d'obtenir une information, mais aussi de la nature de ce processus ; pour reprendre l'analogie informatique, il ne suffit pas de rendre compte du temps que prend l'application d'un algorithme (ou du nombre d'étapes en lesquelles il consiste, à supposer que l'on puisse les compter) mais aussi de la nature de cet algorithme (des règles de transformations elles-mêmes). Comme les différents exemples que j'ai présentés le montrent, des représentations de formats différents non seulement ne rendent pas les informations qu'elles contiennent aussi aisément accessibles (elles n'ont pas le même contenu cognitivement accessible), mais encore elles ne facilitent pas le même *type* d'inférences (on n'accède pas à ces informations de la même manière). Lire et comparer des valeurs sur une liste n'est pas la même opération qu'identifier des couleurs et comparer leurs relations topologiques. Calculer la valeur d'une variable x à un instant t au moyen de la fonction correspondante ne revient pas au même que tracer le graphe de cette fonction et identifier visuellement la valeur de x à t .

Il est certes difficile de donner une définition précise de ces types d'inférences (et, plus largement, d'opérations cognitives). Cependant, sans savoir « ce qui se passe dans la tête » des agents, et sans être capable d'énoncer les algorithmes propres à modéliser ces processus, on a une compréhension intuitive de ce qu'est un type d'inférence,

par analogie avec la notion d'algorithme. Cette compréhension intuitive permet de dire que deux représentations dans des formats différents facilitent différents types d'inférences. Il y a bien un sens dans lequel le processus qui conduit un agent à conclure des tableaux de la figure 32 que la zone nord est plus chaude que la zone sud ouest, en plus d'être plus rapide, n'est pas du même type que celui qui permet d'obtenir la même conclusion à partir de la liste. La définition du format d'une représentation doit non seulement rendre compte du coût, mais aussi du *type* d'opérations cognitives nécessaires pour accéder aux différentes informations qu'elle contient.

2.2.3 Vers une définition de la notion de format

Où en sommes-nous ? L'examen qui précède nous permet de caractériser un changement de format comme un changement dans le coût et le type d'inférences qu'un agent doit effectuer pour accéder aux différentes informations contenues dans une représentation. Alors que le système symbolique (les règles syntaxico-sémantiques) d'une représentation suffisent à en déterminer le contenu informationnel, son contenu cognitivement accessible pour des agents aux capacités cognitives limitées est *aussi* déterminé par les opérations inférentielles effectuées *en pratique* par ces agents pour obtenir différentes informations.

La notion de format doit donc rendre compte non pas des différences *objectives* entre la liste et les tableaux de la figure 32 – dont la notion de système symbolique vise à rendre compte – mais des différences *pour les agents*. Il existe des différences objectives, non seulement entre la liste et les tableaux, mais aussi entre les deux tableaux : leur interprétation n'obéit pas exactement aux mêmes règles syntaxico-sémantiques (le système symbolique du tableau en couleurs comporte « plus » de règles syntaxico-sémantiques que celui du tableau en noir et blanc). Cela, la notion de système symbolique suffit à en rendre compte. Cependant, elle ne rend pas compte de ces différences *pour les agents*. Elle ne permet pas de rendre compte, par exemple, du fait que la même représentation (une carte de température dans des tons verts et rouges) n'est pas dans le même format pour un agent aux capacités perceptives normales et pour un daltonien.

Peut-on, de l'analyse qui précède, tirer une définition de la notion de format ? Comme je l'ai suggéré à plusieurs reprises, mon usage de cette notion repose sur une analogie avec le sens qu'elle a en informatique. Suivant cette analogie, le format d'une représentation devrait pouvoir être défini comme un ensemble de procédures consistant en l'application de règles de construction et de transformation, déterminées par les propriétés syntaxico-sémantiques du système symbolique dans lequel la

représentation fonctionne *et* par les capacités cognitives des agents. Une telle définition permettrait de rendre compte du fait qu'une liste et un tableau ne sont pas dans le même format, mais aussi du fait qu'une même représentation n'est pas dans le même format pour deux agents aux capacités cognitives différentes. Elle rendrait compte, pour chaque agent, du coût et du type d'inférences à mettre en œuvre pour obtenir les différentes informations contenues dans cette représentation.

Cependant, si l'idée d'un ensemble de règles de transformation permettant de traiter les données est une analogie suggestive pour décrire les processus cognitifs des humains, il est douteux que l'on puisse énoncer formellement des algorithmes correspondant à ces processus. Les représentations utilisées par les humains sont en effet beaucoup plus sophistiquées que des suites de symboles numériques binaires, et les processus cognitifs sont beaucoup plus difficiles à connaître que le traitement des données par un ordinateur.⁵¹

En l'absence d'une connaissance détaillée des processus mentaux des agents, la définition des formats de représentation comme des ensembles de règles de construction et de transformation reste donc une analogie. Néanmoins, cette analogie semble nous autoriser à dire que *le format d'une représentation est une fonction du coût et du type d'inférences que des agents aux capacités cognitives limitées doivent effectuer pour accéder aux différentes informations contenues dans cette représentation*. Une représentation joue un certain rôle inférentiel pour un agent. Ce rôle inférentiel peut être défini par toutes les inférences que cette représentation permet aux agents de tirer. Selon les cas, ces inférences sont, on l'a vu, plus ou moins coûteuses, et de types différents. Le format d'une représentation est ce qui détermine la nature et le coût de ces inférences – et donc le rôle inférentiel d'une représentation.⁵²

⁵¹Des hypothèses de différents niveaux ont été proposées sur les processus mentaux des agents pour expliquer le fait qu'ils ont plus de facilité à raisonner avec certains types de représentation. Certains travaux en intelligence artificielle cherchent à modéliser les processus de transformation opérés par les agents sans prétendre que leurs modèles correspondent effectivement aux processus mentaux. L'hypothèse des modèles mentaux, avancée par Johnson-Laird (1983), suggère que les opérations mentales se font sur la base de représentations qui ont elles-mêmes un format bidimensionnel, et non séquentiel comme le suggère l'hypothèse d'une logique mentale. Cette hypothèse se situe au niveau psychologique, et ne prétend pas, chez Johnson-Laird en tous cas, décrire les processus cérébraux qui implémentent les raisonnements. L'hypothèse des modèles mentaux suggère que les représentations spatiales, outre le fait qu'elles permettent, en vertu de leurs propriétés syntaxico-sémantiques, de présenter certaines informations de manière immédiatement accessibles, facilitent les raisonnements en raison de leur parenté avec les représentations mentales : les agents n'ont pas besoin de les « transformer », comme c'est le cas à partir de représentations linguistiques.

⁵²Si je distingue le format d'une représentation de son rôle inférentiel, c'est parce que la notion de format doit permettre de dire que deux représentations qui ont des cibles différentes (et donc aussi des propriétés perceptibles différentes, sauf si leurs cibles sont identiques du point de vue des

Le rôle inférentiel d'une représentation dépend à la fois des propriétés syntaxico-sémantiques de la représentation et des capacités des agents. On peut imaginer qu'une connaissance détaillée des processus mentaux des agents permettrait de définir plus précisément ce rôle ; on pourrait alors définir *le format d'une représentation comme un ensemble de procédures que les agents doivent mettre en œuvre pour accéder aux différentes informations contenues dans cette représentation*.

Dans la prochaine section, je voudrais montrer qu'une telle définition est intenable. Comme une analyse plus attentive de l'exemple des diagrammes de Feynman et de celui des équations de la mécanique le mettra en évidence, il n'est pas seulement impossible en pratique d'énoncer ces règles, mais encore l'idée même d'un ensemble de règles ou d'opérations à suivre, qui serait déterminé par les propriétés syntaxico-sémantiques d'une représentation et les capacités des agents, repose sur une idéalisation grossière. Elle ne permet pas de rendre compte de la complexité de l'interaction cognitive des agents avec les représentations externes utilisées dans la pratique scientifique.

3 Les formats ne sont pas réductibles à des ensembles de règles

On a vu, dans ce qui précède, que le format d'une représentation, qui détermine les processus inférentiels que les agents doivent effectuer afin d'en extraire telle ou telle information, dépend, entre autres, de leurs capacités cognitives : une liste de données numériques n'est pas dans le même format qu'un graphe présentant les mêmes données parce que les agents, en raison de leurs capacités cognitives limitées, ne doivent pas mettre pas en œuvre les mêmes processus inférentiels selon qu'ils utilisent la liste ou le graphe. Ainsi énoncée, la définition du format d'une représentation est normative (elle est définie par le coût cognitif et le type d'inférences que les agents *doivent*

propriétés représentées) peuvent être dans le même format. Par exemple, on doit pouvoir dire qu'une carte de France et une carte d'Angleterre, exactement à la même échelle, contenant le même *type* d'informations (par exemple, le relief, et rien d'autre) à propos de leur cible, sont dans le même format. Elles n'ont pas le même contenu informationnel (et par conséquent pas non plus le même contenu cognitivement accessible). Elles ne jouent donc pas le même rôle inférentiel, puisque les informations qu'elles permettent de tirer ne sont pas les mêmes, mais leur utilisation implique la mise en œuvre de processus inférentiels similaires (en nature et en coût). Elles ont donc le même *type* de rôle inférentiel ; c'est ce qui permet de dire qu'elles sont dans le même format. Pour résumer : le format *et* les propriétés perceptibles d'une représentation en déterminent ensemble le rôle inférentiel. Ce dernier est l'ensemble des inférences que cette représentation permet aux agents de tirer. En conséquence, une différence de format implique une différence de rôle inférentiel, mais l'inverse n'est pas vrai, sauf dans le cas de deux représentations ayant le même contenu informationnel.

mettre en œuvre) ; en conséquence, elle repose sur l'hypothèse d'un agent type. Cet agent, doté de capacités cognitives normales (limitées), est supposé « maîtriser le système symbolique » dans lequel la représentation doit être interprétée. Pour prendre un exemple trivial, affirmer que deux équations différentielles dans des systèmes de coordonnées différents ont des formats différents n'a de sens que si l'on suppose un agent qui maîtrise le calcul différentiel.

Dans cette section, je vais montrer que l'hypothèse d'un agent type qui « maîtrise le système symbolique » d'une représentation repose sur une idéalisation qui devient particulièrement gênante dès que l'on s'intéresse aux représentations sophistiquées que sont les représentations théoriques. À partir du moment où les agents entrent dans la définition du format des représentations, on ne peut se contenter de parler, en général, de leurs « capacités cognitives » et de leur « maîtrise du système symbolique » sans entrer dans plus de détail. En conséquence, comme on va le voir, la définition du format d'une représentation ne peut pas être normative.

3.1 L'expertise comme maîtrise d'un système symbolique ?

J'ai affirmé à plusieurs reprises que l'obtention d'une information à partir d'une représentation suppose que l'on maîtrise le système symbolique dans lequel elle est censée être interprétée. Extraire une information d'un graphe, par exemple, nécessite que l'on maîtrise les règles de construction, de transformation et d'interprétation d'une telle représentation.

Les représentations utilisées dans la pratique scientifique, comme, par exemple, les équations différentielles utilisées en mécanique, supposent, pour qu'on soit en mesure d'en extraire des informations, un certain degré d'expertise. Alors que la lecture d'un plan de métro est accessible à la plupart des adultes citadins dotés de capacités cognitives normales, ce qui autorise à supposer que son rôle inférentiel est, à peu de choses près, le même pour tous, il semble que le rôle inférentiel des équations lagrangiennes soit fort différent pour la majorité des humains d'une part et les physiciens et mathématiciens de l'autre. Il est, de plus, différent pour un débutant et un expert : à supposer que le débutant soit capable de tirer les mêmes conséquences de ces équations, c'est au prix de gros efforts : en conséquence, le coût cognitif qu'il doit fournir est plus important, et le rôle inférentiel des équations n'est pas le même que pour l'expert.

Cela, cependant, ne remet pas directement en cause la définition des formats de représentation esquissée précédemment. De même que le format d'un plan de métro ne prend pas en compte ce qu'un nouveau né ou un aveugle peuvent en faire, il suffit de considérer que la définition du format des représentations utilisées dans la

pratique scientifique est relative à leurs *utilisateurs standard*, c'est-à-dire aux experts. L'existence d'agents incompetents n'est donc pas une objection à cette définition : ces agents, ne maîtrisant pas le système symbolique, ne font pas partie des utilisateurs standard. Ainsi, la définition esquissée dans la section précédente, si elle n'est pas relative à un agent idéal sans limitations cognitives (la notion de format n'a pas de sens pour un tel agent), n'est pas non plus relative à un profane. Elle détermine le rôle inférentiel d'une représentation pour un agent aux capacités cognitives limitées, qui maîtrise les règles de ce type de représentation.

On pourrait même suggérer que l'expertise dans un domaine peut être définie, au moins partiellement, par la maîtrise des règles d'interprétation de certains types de représentations. Devenir expert consisterait alors à acquérir progressivement la maîtrise de ces règles. La définition du format d'une représentation offrirait ainsi une mesure de l'expertise relative des agents.

Quoique cette définition de l'expertise semble assez bien correspondre à l'idée intuitive que l'on s'en fait, je vais montrer que l'hypothèse d'un tel utilisateur standard repose sur une idéalisation qui fait manquer des aspects essentiels de l'utilisation des représentations théoriques par les agents. Ce constat me conduira à renoncer à toute définition normative de la notion de format.

3.2 Qui maîtrise le système symbolique des diagrammes de Feynman ?

Considérons de nouveau le cas des diagrammes de Feynman. À la section précédente, j'ai insisté sur la différence de format *entre les formules mathématiques et les diagrammes*, et j'ai montré que, bien qu'équivalents en principe, ces deux types de représentations ne permettent pas les mêmes calculs et les mêmes développements théoriques.

L'intérêt de cet exemple pour mon propos n'est cependant pas épuisé par la comparaison de ces deux types de représentation. L'utilisation des diagrammes eux-mêmes par différents physiciens – Feynman, Dyson, mais aussi d'autres chercheurs appartenant à des « écoles » et travaillant dans des contextes différents – nous offre un exemple particulièrement net du fait que des représentations théoriques peuvent avoir un rôle inférentiel différent pour différents experts. Je me contenterai, dans ce qui suit, de rapporter quelques éléments de l'enquête historique détaillée de David Kaiser (2005) sur ce qu'il appelle la « propagation » des diagrammes de Feynman au sein de la communauté des physiciens théoriciens de l'après-guerre, et de renvoyer à son travail pour plus de détails. Je me concentrerai, en particulier, sur ce qu'il appelle la « scission entre Feynman et Dyson ».

Feynman et Dyson, dans leur correspondance et leurs écrits personnels, expriment en effet à plusieurs reprises un désaccord explicite sur l'usage qu'ils jugent légitime de faire des diagrammes et sur leur statut au sein du cadre théorique de l'électrodynamique quantique. Dyson, selon Kaiser, les considère comme des *représentations dérivées des formules* mathématiques, destinées à remplir la fonction d'aides psychologiques pour effectuer des calculs difficiles. Il affirme que l'usage des diagrammes serait illégitime s'ils n'étaient pas rigoureusement dérivables des formules mathématiques : « jusqu'à ce que ses règles soient codifiées et rendues précises d'un point de vue mathématique, je ne pouvais pas appeler [la méthode de Feynman] une théorie. » (Dyson, 1951, p. 127) Pour lui, les diagrammes sont des moyens de « visualiser les formules [qu'il dérive] rigoureusement de la théorie des champs. » (pp. 129-130) ; les diagrammes n'ont de signification qu'*au sein* de l'électrodynamique quantique, à laquelle ils n'apportent rien d'autre qu'une facilité de calcul.

En revanche, Feynman affirme, comme le rapporte Kaiser, n'avoir jamais ressenti le besoin de montrer comment les diagrammes peuvent être obtenus à partir des formules mathématiques. Il exprime, à plusieurs reprises, sa *préférence* pour les diagrammes : « Toutes les preuves mathématiques furent des découvertes ultérieures que je ne comprends pas tout à fait, mais les idées physiques me semblent très simples. » (Lettre à Ted Welton, 16 novembre 1949, citée dans Kaiser, 2005, p. 176)⁵³ Contrairement à Dyson, il considère donc les diagrammes comme *premiers* et plus importants que toutes les dérivations mathématiques que l'on peut en tirer. En plus d'être des aide-mémoire, les diagrammes fournissent, selon lui, une dimension intuitive à la théorie : il les considère comme des « images intuitives » (Kaiser, 2005, p. 176). Comme le remarque Dyson, Feynman « consid[ère] le graphe [le diagramme] comme l'image d'un processus se produisant effectivement dans l'espace-temps » (Dyson, 1951, p. 127). Plutôt que des visualisations *des formules*, les diagrammes sont pour Feynman des visualisations *des processus physiques* eux-mêmes.

Ainsi, malgré leur accord sur l'équivalence en principe des diagrammes et des formules, Feynman et Dyson ne font pas le même usage des diagrammes. En premier lieu, ils ne s'entendent pas sur le statut qu'il convient d'accorder à des représentations diagrammatiques dans un domaine où le « griffonnage » n'a, traditionnellement, qu'une valeur d'aide psychologique. Autrement dit, alors que Dyson considère que les diagrammes ne font pas partie de l'appareil théorique de l'électrodynamique quantique, Feynman les considère comme des représentations théoriques au même titre que les équations mathématiques.

⁵³Feynman parle également de la « plausibilité physique » de l'approche diagrammatique (Kaiser, 2005, p. 177).

Ce désaccord ne se réduit pas, cependant, à la confrontation de différents jugements de valeurs, ou même de deux conceptions de la physique théorique. Affirmer que les diagrammes font partie de l'appareil théorique au même titre que les équations revient à dire qu'ils ne tirent pas leur signification des formules mathématiques ; bien au contraire, pour Feynman, ce sont les formules mathématiques que l'on peut retrouver à partir des diagrammes qui tirent leur signification de ces derniers. Ainsi, alors que Feynman trace « intuitivement » ses diagrammes, Dyson les tire rigoureusement des formules mathématiques. Pour le premier, ils représentent les processus physiques eux-mêmes ; pour le second, ils ne font que présenter sous une forme spatiale le contenu des formules mathématiques. Ne construisant et n'utilisant pas les diagrammes de la même manière, Feynman et Dyson, finalement, n'y « voient » pas la même chose.

Comme le suggère Kaiser, cette différence dans l'utilisation des diagrammes par les deux physiciens trouve des éléments d'explication dans leurs propres engagements et préférences théoriques. Contrairement à Dyson, qui s'applique à démontrer comment les diagrammes de Feynman et les équations de Schwinger peuvent s'insérer dans un cadre cohérent de la théorie des champs⁵⁴, la méthode de renormalisation de Feynman, au sein de laquelle les diagrammes ont vu le jour, est fondée sur une approche en termes de particules, et non de champs. Plus généralement, Kaiser note que Feynman a une préférence pour une approche semi-classique, et travaille presque toujours en termes de particules, essayant de retirer entièrement les champs des descriptions théoriques. Ces engagements et ces théoriques, joints aux préférences individuelles pour certains types de raisonnement (Feynman revendique à plusieurs reprises son penchant pour la « visualisation », par rapport au calcul abstrait) doivent avoir contribué à donner aux diagrammes des rôles inférentiels différents pour Feynman et pour Dyson. Comme je l'ai dit, Dyson déduit les diagrammes des formules mathématiques, quand Feynman les dessine intuitivement : chacun les relie différemment aux autres représentations et, finalement, au monde physique.

S'il était possible de décrire l'utilisation que Feynman et Dyson font chacun des diagrammes au moyen de règles de construction et de transformation, il est clair que l'on en conclurait que ceux-ci ne sont pas dans le même format pour les deux physiciens. Mais ce n'est pas tout. Feynman lui-même, comme le rapporte Kaiser (2005), et en même temps que lui d'autres physiciens, travaillant dans des groupes différents (Oxford et Cambridge, Japon, Union Soviétique), ont continuellement modifié leurs règles de construction et d'utilisation des diagrammes. L'analyse que Kaiser (2005)

⁵⁴Dyson (1951, p. 23) prétend avoir contribué à permettre à « des gens comme Pauli, qui cro[ient] dans la théorie des champs, de dessiner des diagrammes de Feynman sans abandonner leurs principes ».

propose de la « plasticité » des diagrammes au cours de leur « propagation » dans les pratiques de la physique théorique moderne lui permet de montrer que, selon les contextes, leur utilisation et leur interprétation – leurs « modes d’emplois » – varient.

Malgré les efforts de Dyson, les règles de construction et d’interprétation des diagrammes n’ont pas été appliquées à la lettre, loin de là. C’est, de plus, une des raisons de leur succès : plutôt que de simples outils de calcul, ils sont devenus d’authentiques outils de découverte et d’invention théoriques, qui ont contribué de manière essentielle à d’importants développements de la physique moderne, en étant appliqués à d’autres domaines.

Quel est donc l’« utilisateur standard » de ces diagrammes ? Feynman ? Dyson ? Les membres de telle ou telle école ? Quel est l’expert qui maîtrise le système définissant le format des diagrammes et en fixant le contenu ? Le rôle inférentiel des diagrammes est manifestement différent pour Feynman et pour Dyson. De plus, ce rôle change constamment pour Feynman lui-même. Réduire l’utilisation des diagrammes à un ensemble de règles que leurs utilisateurs *doivent* maîtriser et auxquelles ils doivent se conformer revient à manquer l’essentiel de leur rôle dans l’activité théorique. En conséquence, il semble que la définition du format d’une représentation proposée à la section précédente ne soit pas tenable, tout au moins pour le cas des diagrammes de Feynman. Je voudrais à présent suggérer que, loin d’être une exception, il s’agit là d’un cas exemplaire de la manière dont les représentations sont utilisées dans la théorisation.

3.3 Les équations de la mécanique : le bon algorithme n’existe pas

L’exemple des diagrammes de Feynman, dont je n’ai fait que rapporter quelques éléments empruntés à l’étude de Kaiser (2005), peut sembler, de prime abord, tout à fait exceptionnel. Pour nous convaincre du contraire, revenons un instant à l’exemple de la mécanique classique.

Il peut sembler que, dans le cas des équations newtoniennes, lagrangiennes et hamiltoniennes, il existe des ensembles de règles bien définies que tout expert maîtrise. La procédure permettant de résoudre un problème de mécanique au moyen de ces différentes équations, semble en effet entièrement dictée par un protocole précis. Si, comme je l’ai suggéré dans la section précédente, on peut affirmer que ces trois types d’équations sont dans des formats différents, il semble qu’il s’agit là d’une différence objective, au sens où elle ne dépend pas des agents ; rien ne permet de dire que chacune de ces équations peut être utilisée de diverses façons, du moins par des agents compétents.

Admettons que ce soit le cas, et que tous les physiciens, aujourd’hui, les utilisent

en suivant les mêmes processus inférentiels. Peut-on affirmer, pour autant, que les équations de Lagrange ont le même rôle inférentiel avant et après l'introduction des transformations de Legendre par Hamilton ? Comme je l'ai dit au chapitre 3, cette innovation ne repose pas sur une nouveauté empirique, qui aurait pu conduire Hamilton à modifier certaines hypothèses théoriques, mais consiste en l'introduction de règles de transformations mathématiques permettant, à partir d'une équation du second ordre, d'obtenir deux équations du premier ordre. Autrement dit, cette innovation consiste en un changement de format.

Pour un agent aux capacités cognitives illimitées, ce développement formel ne changerait rien : le contenu informationnel des équations de Lagrange ne s'en trouve aucunement modifié. Pourtant, introduire les transformations de Legendre en mécanique permet de transformer des équations insolubles en équations solubles. Non seulement les équations hamiltoniennes ne sont pas dans le même format que les équations lagrangiennes, mais il est un sens dans lequel on peut dire que les équations lagrangiennes elles-mêmes ne sont pas dans le même format pour quelqu'un qui connaît les transformations de Legendre et pour quelqu'un qui les ignore : ces règles de transformation, en s'ajoutant au « mode d'emploi » des équations lagrangiennes, en modifient le contenu cognitivement accessible. En ce sens, les travaux de Hamilton ont bien contribué à modifier le contenu des différents principes et équations de la mécanique.

Le développement de la mécanique classique, dont on a vu que, du point de vue d'une conception logico-empirique des théories, il consiste en un développement « purement formel », peut ainsi être décrit comme un processus de modification du rôle inférentiel des différents principes et équations. Comme un joueur d'échec qui, connaissant les règles du jeu, et s'y conformant parfaitement, invente de nouveaux coups, le théoricien modifie les processus inférentiels qu'il suit dans le but de résoudre des problèmes et, parfois, de développer de nouveaux liens entre différents principes de sa théorie. Autrement dit, il ne se conforme pas à un format – un ensemble de procédures à mettre en œuvre – mais il en invente de nouveaux.

Dans le cas de la mécanique, tout comme dans celui de Feynman, définir le format des différentes équations comme un ensemble de règles qu'un agent *doit* suivre dans le but d'obtenir des informations repose sur une idéalisation illégitime : le bon algorithme n'existe pas. En conséquence, vouloir ainsi réduire les représentations théoriques à des représentations dont le rôle inférentiel est déterminé de manière fixe par un ensemble de procédures (le supposé format) reviendrait à manquer des aspects essentiels de la découverte et de l'invention scientifiques, que l'on peut décrire, dans certains cas, comme l'*invention de nouveaux chemins inférentiels* entre les différents principes d'une théorie.

3.4 Renoncer à une définition normative des formats

Ces considérations suggèrent que l'on doit renoncer à l'idée d'une définition normative des formats. Ce que l'on appelle intuitivement « format » doit plutôt recevoir une définition descriptive : le format d'une représentation *est* le rôle inférentiel⁵⁵ qu'elle joue *effectivement* pour un agent particulier, dans une *situation* particulière. Autrement dit, ce n'est pas l'ensemble des procédures que cet agent doit mettre en œuvre, mais celles qu'il met effectivement en œuvre. Ces procédures dépendent, en plus des propriétés perceptibles de la représentation et d'un ensemble minimal de règles de construction et d'interprétation, des capacités cognitives des agents, ainsi que, comme le cas des diagrammes de Feynman le met en évidence, de ses engagements théoriques, de ses préférences, et de ses habitudes de raisonnement.

Il convient de préciser que cela ne revient pas à dire que le raisonnement scientifique n'obéit à aucune règle et qu'il n'existe pas de différence entre une inférence correcte et une inférence fautive. Bien entendu, résoudre une équation différentielle implique de se conformer à tout un ensemble de règles de calcul. Si on les suit, on ne peut pas tirer de conclusions contradictoires des différentes équations de la mécanique. Le rôle inférentiel d'une représentation est partiellement déterminé par tout un ensemble de règles objectives, qui ne dépendent pas de l'utilisateur et de la situation. Cependant, *ces règles sous-déterminent les processus inférentiels des agents*, tout comme les règles du jeu d'échec, aussi contraignantes soient-elles, sous-déterminent la trajectoire que les joueurs vont faire emprunter à leurs pièces.

Les représentations théoriques sont des outils hautement sophistiqués de calcul et d'enquête, dont les modes d'emplois ne sont pas établis d'avance. L'activité théorique consiste en partie à développer ces règles d'utilisation pour explorer le contenu des représentations. L'idée selon laquelle deux équations différentes ont le même contenu informationnel, tout comme l'idée selon laquelle une même équation a le même contenu informationnel pour deux experts différents, ne rend pas compte de la manière dont les représentations sont utilisées et comprises par les agents. L'analyse de la notion de format permet ainsi de mettre en évidence la différence fondamentale entre une approche logico-empirique du contenu des théories *en principe* et celle que je défends dans l'ensemble de ce travail, qui consiste à placer au premier plan ce que les agents, *en pratique*, en font. Dans la prochaine section, je propose de souligner le lien entre mon examen de la notion de format et la définition de la notion de version, et d'en tirer quelques conséquences pour une analyse de l'activité théorique.

⁵⁵Ou, plus précisément, le *type* de rôle inférentiel (voir note 52 page 406).

3.5 Versions, formats et activité théorique

J'ai défini, au chapitre 3, la version d'un agent comme l'ensemble des représentations mentales qu'il se forme et utilise au cours de son apprentissage et de son utilisation d'une théorie – c'est-à-dire d'un ensemble d'hypothèses théoriques. L'hypothèse selon laquelle chaque agent, du profane à l'expert, a sa propre version d'une théorie est destinée, comme je l'ai alors signalé, à rendre compte de l'activité théorique, en tant que cette expression désigne aussi bien l'activité de l'étudiant qui apprend, que celle du théoricien qui développe et applique des hypothèses. J'ai en effet proposé de concevoir l'activité théorique comme consistant (entre autres) en un processus de développement et de renforcement de cet ensemble de représentations mentales. L'étudiant qui apprend, comme l'expert qui approfondit un ensemble d'hypothèses théoriques, développe et renforce les liens déductifs entre ces hypothèses et les rend applicables à des cas de plus en plus nombreux et divers. Cette description de l'activité théorique, comme je l'ai plusieurs fois souligné, vise à montrer que, selon la manière dont les agents l'organisent et le relient au monde, un ensemble d'hypothèses théoriques (une « théorie ») n'a pas exactement le même contenu.

L'analyse que je viens de proposer de la notion de format permet de rendre compte de façon plus précise de l'activité théorique ainsi définie. Elle permet en particulier d'approfondir et de renforcer l'hypothèse selon laquelle le contenu d'une représentation est indissociable non seulement de sa forme, mais aussi de la manière dont chaque agent, selon ses particularités cognitives, ses compétences, ses engagements et ses intérêts théoriques, utilise cette représentation, présentée sous une forme particulière. On comprend en effet que chaque agent, tout en respectant les règles objectives qui déterminent la correction des inférences que l'on peut tirer au moyen d'une représentation, emprunte différents chemins inférentiels et dessine sa propre carte mentale (sa propre version) qu'il relie au monde de différentes manières.

Cette analyse permet ainsi, comme je l'ai suggéré, de souligner la parenté entre les processus d'apprentissage, d'application et de développement d'un ensemble d'hypothèses théoriques. Si l'apprentissage d'une théorie ne débouche pas sur le développement de cette dernière, au sens où l'étudiant n'invente (en général) pas de liens déductifs inédits, on peut décrire ce processus comme celui de la modification progressive du rôle inférentiel de certaines représentations chez le débutant. Alors que, comme je l'ai souligné au chapitre 3, le mot « force » a seulement, pour un novice, le sens qu'il a dans le langage ordinaire – lui-même déjà loin d'être univoque –, apprendre la mécanique consiste en partie à modifier le rôle de ce mot dans ses raisonnements, ce qui passe à la fois par l'acquisition de compétences computationnelles (en particulier celles du calcul différentiel) et de compétences représentationnelles (il

s'agit d'apprendre comment relier ce mot au monde empirique).

L'expertise peut bien, comme je l'ai envisagé précédemment (page 408), être décrite comme la maîtrise de certains systèmes de représentation, à ceci près que cette maîtrise ne se mesure pas (seulement) à l'aune de règles préexistantes auxquelles l'utilisateur doit se conformer. Comme le suggère Andrea Woody (2004), devenir expert dans un domaine consiste à acquérir une « conscience articulée » des représentations utilisées dans ce domaine. Cette « conscience articulée » peut être indéfiniment développée : il est toujours possible, tout en se conformant aux inférences permises par les règles objectives d'un type de représentation, d'inventer de nouvelles règles de transformation, qui modifient le rôle inférentiel de ces représentations. Comme on l'a vu, personne ne maîtrise « parfaitement » l'usage d'un type de représentation comme, par exemple, celui des diagrammes de Feynman.

Ainsi, approfondir sa compréhension d'une théorie – développer sa version – consiste à développer de nouveaux chemins inférentiels entre différentes représentations et à réduire le coût inférentiel que l'on aura typiquement à payer pour tirer l'une de l'autre ; cela peut déboucher, dans certains cas, sur l'invention d'un nouveau type de représentation, qui ouvre des perspectives d'applications nouvelles. C'est ce que les étudiants en physique apprennent. C'est ce que Hamilton a fait, à cette importante différence près qu'il l'a fait en premier, et qu'il a rendu sa modification publique.

Pour finir, notons que cette analyse permet de renforcer un certain nombre de suggestions que j'ai tirées, au chapitre 6, de l'examen des propos de Kuhn. Au cours de cet examen, j'ai distingué entre deux conceptions de la signification qui sous-tendent tour à tour les différentes thèses de Kuhn. J'ai montré, en particulier, que ses propos sur l'incommensurabilité reposent sur une conception holiste et vériconditionnelle de la signification, alors que, à d'autres moments, Kuhn semble défendre l'idée selon laquelle la signification d'une représentation dépend de la manière dont les agents l'utilisent pour résoudre des problèmes. C'est cette seconde conception, dont mon analyse est un approfondissement, qui permet à Kuhn de dire que l'équation de Schrödinger (et, par extension, la mécanique quantique) n'a pas exactement la même signification pour des groupes d'agents qui l'utilisent dans des cas paradigmatiques différents (voir chapitre 6, page 291).

L'analyse de la notion de format, et en particulier l'exemple des diagrammes de Feynman, corrobore cette conception de la signification, en montrant qu'elle semble bien adaptée au cas des représentations théoriques, dont la nature et même l'existence des référents dans le monde empirique est loin d'aller de soi. Deux utilisateurs peuvent bien s'accorder sur un ensemble de règles de manipulation d'une représentation sans pour autant en tirer les mêmes inférences et la relier de la même manière au monde

empirique. Comme on l'a vu, les diagrammes de Feynman ne sont pour Dyson que des visualisations des équations, lesquelles sont rattachées au monde empirique par l'attribution d'une signification aux symboles représentant des champs, alors que, pour Feynman, ce sont d'authentiques représentations des phénomènes, dans une perspective semi-classique, qui exclut la notion de champ (voir page 410). De même, l'expression « $f = ma$ » ne se voit pas attribuer exactement la même signification et n'est pas rattachée au monde empirique de la même manière par Mach, qui préfère se représenter les phénomènes en termes de forces (voir chapitre 3, section 4.1.1) et par Duhem, qui préfère parler d'énergie (voir chapitre 3, section 4.3, en particulier page 147).

Cela ne doit pas, pour autant, faire conclure à des cas d'incommensurabilité totale et d'échec de la communication. Comme je l'ai souligné au chapitre 6 (page 314), la conception inférentialiste de la signification qui sous-tend mon analyse autorise à envisager des cas de similitude partielle de signification. Les différentes manières d'utiliser les diagrammes de Feynman et les équations de la mécanique sont des façons valables de les comprendre, au sens où les inférences qu'elles mettent en jeu respectent un ensemble de règles objectives. De même, les versions des différents acteurs de l'histoire de la mécanique, exprimées dans les présentations étudiées au chapitre 3, correspondent à leur utilisation des différents principes dans leurs raisonnements et leur application : la manière dont ils l'ont apprise, leurs préférences, leurs engagements théoriques d'arrière-plan, déterminent le rôle inférentiel des principes et des équations pour chacun d'entre eux, sans que cela remette en cause la possibilité de la communication. Que Mach et Duhem ne comprennent pas « $f = ma$ » de la même manière n'implique cependant pas que leurs versions soient incommensurables. Le fait que chaque individu, à la rigueur, ait sa propre carte mentale des différentes représentations – sa propre version, par exemple, de la mécanique – n'implique pas qu'il ne puisse pas s'accorder avec les autres sur certaines – et mêmes la plupart – des utilisations de ces représentations.

Comme je l'ai suggéré au chapitre précédent, en accordant de l'importance à la place particulière qu'occupe une représentation dans l'ensemble des représentations mentales d'un agent, mon analyse permet de défendre une sorte de holisme *situé*, qui se distingue du holisme global que son adoption d'une conception vériconditionnelle de la signification conduit Kuhn à défendre. L'ancrage des notions de format et de version dans une conception inférentielle de la signification permet ainsi de renouer avec l'idée d'une théorie comme ensemble de représentations, à ceci près qu'elle n'est plus conçue comme un tout représentationnel possédant un certain contenu déterminé par des règles d'application, mais plutôt comme un réseau de représentations permettant de tirer certaines inférences avec plus ou moins de peine, et relié par

différents « bouts » au monde empirique.

Conclusion

Le constat de l'importance de la forme des représentations utilisées par les agents dans la pratique scientifique m'a conduit à chercher des outils d'analyse permettant de rendre compte, d'une part, de la diversité des types de représentations et, d'autre part, des conséquences de cette diversité pour les processus cognitifs des agents. La notion de système symbolique proposée par Goodman m'a permis d'esquisser une typologie de certaines représentations fréquemment utilisées en science, fondée sur la relation entre les propriétés perceptibles de ces représentations et leur contenu.

L'analyse de la notion de format de représentation, initialement motivée par le constat de l'importance de la *forme* particulière des représentations, m'a conduit à montrer que cette forme doit être étudiée dans son interaction avec les agents qui l'utilisent dans une situation particulière. En effet, dès lors que l'on distingue le contenu informationnel d'une représentation de son contenu accessible à des agents aux capacités cognitives limitées, définir ce contenu implique de prendre en compte non seulement la forme de la représentation, mais également les capacités cognitives de l'agent qui l'utilise, ainsi que ses intérêts épistémiques dans une enquête particulière. Comme on l'a vu, non seulement les diagrammes de Feynman ne sont pas dans le même format que les formules mathématiques de Schwinger, mais leur format n'est pas le même pour Feynman et pour Dyson ; il n'est pas non plus le même pour Feynman en 1948 et dix ans plus tard.

Cette perspective permet de rendre compte de plusieurs aspects importants de l'apprentissage et de la recherche scientifiques. La troisième partie de ce travail vise à ainsi à mettre en œuvre les outils élaborés ici dans une nouvelle étude de cas.

Conclusion de la deuxième partie

Après avoir montré, au chapitre 5, les limites internes du projet de reconstruction formelle des théories et, ce faisant, justifié l'entreprise d'une étude de l'activité théorique du point de vue des agents, j'ai cherché, au chapitre 6, à préciser les questions qu'une telle étude doit aborder.

Par l'examen successif des notions de paradigme, de modèle et de gabarit, j'ai ainsi montré que l'analyse de l'interaction cognitive des agents avec les représentations qu'ils utilisent dans la pratique scientifique devrait permettre d'éclairer l'articulation des deux compétences mises en œuvre dans toute activité théorique : celle, que j'appelle computationnelle, qui consiste à manipuler des « morceaux de théorie » sous une forme particulière, et celle, que j'appelle représentationnelle, qui consiste à les utiliser pour représenter des situations particulières.

Dans le chapitre 7, j'ai proposé des outils d'analyse permettant de rendre compte de l'utilisation de représentations de différents types dans la pratique scientifique. D'une part, j'ai suggéré que certains types de représentations sont plus adaptés que d'autres à l'expression de certains types de contenus. D'autre part, j'ai montré qu'une analyse de la relation entre la forme particulière des représentations et les capacités cognitives et intérêts épistémiques des agents invite à renoncer à l'idée selon laquelle le contenu des représentations théoriques est entièrement déterminé par un ensemble de règles objectives. L'analyse de la notion de format m'a ainsi permis de montrer la richesse et la diversité des inférences permises par différents types de représentations théoriques.

Au cours de cette analyse, j'ai montré qu'un ensemble d'hypothèses, que l'on appelle habituellement « théorie », peut être compris et utilisé de plusieurs manières différentes par les agents, sans que cela implique contradiction ou incommensurabilité. Souligner la diversité des versions d'une même théorie permet en outre d'atténuer les frontières *entre* différentes théories.

Dans la troisième et dernière partie de ce travail, je souhaite montrer, par une étude de cas en génétique classique, que l'analyse des types de représentations utilisées dans un domaine scientifique permet d'éclairer certaines questions conceptuelles

propres à ce domaine. Je distinguerai en effet, au sein de ce que l'on appelle « génétique classique », plusieurs composantes théoriques différentes, dont je montrerai qu'elles sont exprimées par des représentations de type différents. L'analyse de la manière dont les généticiens, en fonction de leurs engagements théoriques, utilisent et comprennent ces représentations, permet d'apporter un éclairage nouveau sur la nature et le statut des différentes hypothèses qui constituent la génétique classique. Comme je le montrerai, cette perspective permet également d'aborder la question des relations-interthéoriques de manière nouvelle.

Troisième partie

Représentations en génétique classique

Vue d'ensemble de la troisième partie

Les deux premières parties de ce travail ont cherché à établir la thèse selon laquelle la notion classique de théorie n'est pas une unité d'analyse bien adaptée à l'étude de l'activité théorique. Ce constat, accompagné du parti pris consistant à étudier l'activité théorique du point de vue des agents, m'a conduit à prêter attention aux représentations particulières utilisées dans l'apprentissage et la pratique scientifiques. J'ai supposé que chaque agent, selon l'apprentissage qu'il a suivi, les autres connaissances théoriques dont il dispose, ses engagements et ses talents individuels, se forme sa propre version de l'ensemble des hypothèses auxquelles il est confronté. Cette version se manifeste, entre autres, par la manière dont il utilise les représentations externes dont il dispose.

Tout au long de ces deux parties, mon exemple principal était celui de la mécanique classique. C'est à dessein que j'ai choisi, pour montrer les limites de la notion de théorie, l'exemple canonique par excellence de ce que l'on appelle couramment « théorie ». Je voudrais à présent me tourner vers un domaine, celui des sciences biologiques, où les « théories » sont plus difficiles à identifier et à isoler, et où l'inadéquation des conceptions classiques de la connaissance scientifique est plus patente – et souvent soulignée – que dans le cas des sciences physiques.

Alors que, dans le cas de la mécanique, la diversité des représentations utilisées se réduit, pour l'essentiel, à celle des équations, la biologie est un domaine où l'usage de représentations non linguistiques est central. Je souhaite montrer, dans les deux chapitres qui constituent cette partie, qu'étudier l'activité théorique propre à la génétique classique en prêtant attention aux types de représentations utilisés par les généticiens permet d'éclairer l'articulation des différentes composantes théoriques de ce domaine.

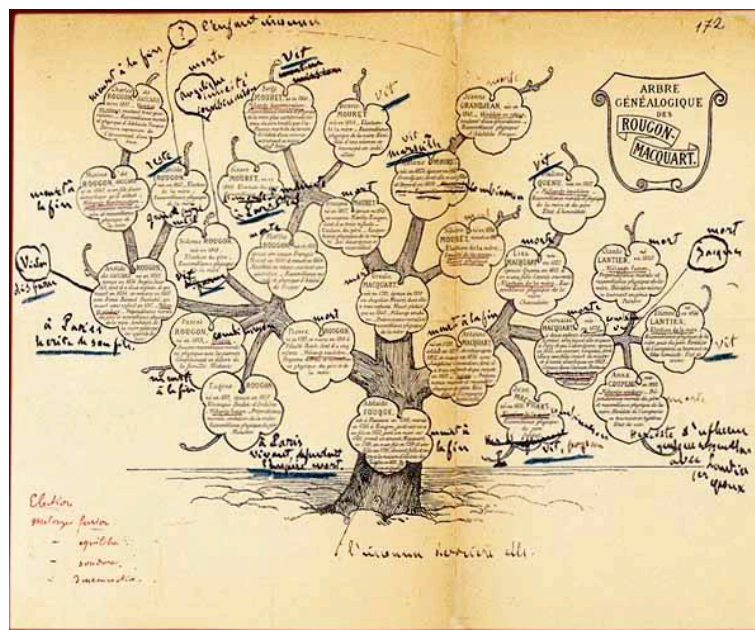
Au chapitre 8, je montrerai que ce que l'on appelle « génétique classique » repose sur l'articulation de deux ensembles d'hypothèses théoriques différents, chacun identifiable à un certain type de pratiques représentationnelles, celui du mendélisme

que j'appellerai « pur », et celui de la cytologie. Je montrerai, par une comparaison avec l'approche de Philip Kitcher (1984), qui répond au même type de motivations que la mienne – une critique des approches formelles des théories – qu'une analyse centrée sur les pratiques représentationnelles plutôt que sur des schèmes de raisonnement abstraits permet d'analyser plus finement l'activité théorique. Je suggérerai en particulier qu'elle permet de jeter un éclairage nouveau sur la question des relations inter-théoriques.

Le chapitre 9 mettra cette suggestion en œuvre, en appliquant l'analyse de la génétique classique proposée au chapitre 8 à l'examen de la construction et de l'interprétation, dans les années 1920, des cartes génétiques, qui occupent une place centrale dans les progrès de la génétique et qui incarnent la rencontre de deux domaines théoriques différents, celui du mendélisme et celui de la cytologie. Cette étude de cas analysera dans le détail la manière dont la construction et l'interprétation des cartes génétiques par différents généticiens est guidée par leurs engagements théoriques, et manifestent leurs versions du mendélisme.

Chapitre 8

Mendélisme et cytologie : de quoi parle la génétique classique ?



Émile Zola, Arbre généalogique des Rougon-Macquart annoté, 1892. Imprimé avec corrections autographes, BnF, Manuscrits.

Ce chapitre et le suivant, consacrés à la génétique classique, sont destinés à mettre en œuvre la démarche suggérée et les outils proposés dans les chapitres précédents. À la différence d'une approche qui consisterait à définir la génétique classique comme une théorie identifiable à une structure hypothético-déductive, je me propose de la caractériser comme un certain type d'*activité théorique* en repérant un ensemble de méthodes et de pratiques représentationnelles typiquement mises en œuvre dans l'analyse classique des problèmes de l'hérédité.

La génétique classique a pour objet l'explication et la prédiction de la transmission héréditaire des caractères chez les êtres vivants. Un des phénomènes dont elle doit typiquement rendre compte est celui de la « réapparition », chez un individu, de certains caractères de ses ancêtres, momentanément « disparus » à la génération précédente : par exemple, le fait qu'un enfant né de deux parents aux yeux bruns puisse avoir les yeux bleus. Pour expliquer ces phénomènes, la génétique classique fait l'hypothèse de l'existence de facteurs héréditaires – les « gènes » – responsables des caractères apparents des individus, mais eux-mêmes inobservables. L'analyse classique de l'hérédité est celle de la transmission de ces facteurs héréditaires, qu'elle décrit au moyen de lois probabilistes ; elle ne dit rien, en revanche, de leur nature ni de leur mode d'action (la manière dont ils causent les caractères).

Cette dernière caractéristique fait de la génétique classique une science hautement « abstraite » – elle est également appelée « génétique formelle » (voir Morange, 2003, p. 26) – au sens où les entités dont elle parle et dont elle décrit le comportement (les gènes) n'ont pas de nature bien définie. Plus exactement, ce sont des unités hypothétiques introduites dans les raisonnements destinés à prédire et expliquer la transmission des caractères, sans que soit faite aucune hypothèse explicite sur leur nature. Malgré la complexité des processus étudiés par la génétique moléculaire, née en 1953 avec la découverte par James Watson et Francis Crick (1953a; 1953b) de la structure en double hélice de l'ADN, elle est, de ce point de vue, beaucoup moins abstraite. Ses concepts, et celui de gène en premier lieu, sont certes difficiles à saisir de manière univoque (voir Waters, 1994; Burian, 1985b, 2005; Kitcher, 1982a; Rheinberger, 2000; Rheinberger et Müller-Wille, 2007) mais il est relativement aisé de spécifier la nature des entités dont elle décrit le comportement : ce sont des molécules, et les phénomènes qu'elle décrit sont des processus chimiques.¹

Née de la redécouverte des travaux de Mendel en 1900, la génétique classique a été principalement développée au cours des premières décennies du vingtième siècle par les travaux de Thomas Morgan et de ses élèves à l'Université de Columbia ; ces der-

¹Pour une analyse des difficultés rencontrées dans l'enseignement de la génétique classique, en raison même de cette abstraction qui la caractérise, voir les réflexions du didacticien Guy Rumelhard (1986).

niers, par l'élaboration de techniques expérimentales et l'introduction de nouveaux types de représentation des phénomènes destinés à rendre compte d'importantes découvertes empiriques, aboutissent dès les années 1920 à la formulation d'un ensemble de concepts qui sont peu ou prou ceux que l'on enseigne aujourd'hui encore sous le nom de « génétique classique ».

À cette période (dès les années 1910), le développement de l'étude mendélienne des phénomènes héréditaires est lié à celui d'une hypothèse au niveau cytologique (c'est-à-dire au niveau des processus cellulaires), celle selon laquelle les chromosomes sont le support matériel des gènes. Cette hypothèse, que l'on appelle couramment la « théorie chromosomique de l'hérédité », est généralement considérée comme une des composantes de la génétique classique.

Historiquement, on peut décrire la théorie chromosomique de l'hérédité comme une première étape vers la *matérialisation* des gènes (voir Jacob, 1970; Gayon, 2000), avant la révolution moléculaire qui ouvre la perspective d'une compréhension de la nature chimique et du mécanisme d'action de gènes : il y aurait ainsi le mendélisme « pur », selon lequel le gène est une unité abstraite, puis la génétique chromosomique qui conçoit le gène comme un segment de chromosome, et enfin la génétique moléculaire, dont on peut dire qu'elle fait du gène une séquence d'ADN, même si c'est une simplification abusive.

Il convient de préciser ce en quoi consiste cette « matérialisation progressive ». Ses étapes successives, en effet, ne peuvent pas simplement être décrites comme des approfondissements et des développements du mendélisme. En particulier, le « passage » du mendélisme pur à la génétique classique et celui de la génétique classique à la génétique moléculaire ne sont pas de même nature. S'il est vrai que la génétique classique, en tant qu'elle inclut l'hypothèse chromosomique, semble faire du gène une unité moins abstraite que le mendélisme pur, elle ne contient cependant pas de définition précise de la nature matérielle des gènes ; les modalités même de la « localisation » des gènes « sur » les chromosomes restent l'objet de spéculations et aucune thèse consensuelle ne se dégage à ce propos. Les recherches des généticiens et des cytogénéticiens dans les années 1930 et 1940, en conduisant progressivement à l'élucidation de la nature chimique des gènes – qui sera l'objet de la génétique moléculaire – ne modifient pas particulièrement les composantes théoriques de la génétique classique, telle que Morgan et ses élèves la comprennent et la pratiquent dans les années 1920, et telle qu'elle est encore enseignée et utilisée aujourd'hui. Autrement dit, la génétique classique, même sous-tendue par l'hypothèse selon laquelle les chromosomes sont le support matériel des gènes, est conceptuellement indépendante des hypothèses sur la nature matérielle des gènes.

Un des buts principaux de ce chapitre est d'analyser l'articulation des différentes

composantes théoriques de la génétique classique – celles du mendélisme pur et celles de la théorie chromosomique. La question de l'identité de la génétique classique a souvent été abordée par les philosophes des sciences ; la plupart du temps, le but d'un tel examen est de répondre à la difficile question des relations entre génétique classique et génétique moléculaire (voir Schaffner, 1969b; Hull, 1972, 1974, 1979; Kitcher, 1984; Waters, 1990). Afin de savoir si l'on peut considérer, comme le suggèrent les biologistes et certains philosophes (en particulier Schaffner, 1969b; Waters, 1990), que les développements de la génétique moléculaire ont opéré une réduction explicative de la génétique classique, il faut au préalable définir ce que l'on entend par « génétique classique ». Ce problème est généralement abordé dans une perspective qui consiste à reconstruire la structure hypothético-déductive des deux théories (génétique classique et génétique moléculaire) afin d'en étudier, ensuite, les relations. Une telle reconstruction contribue à « aplanir » les différentes composantes de la génétique classique et à gommer la distinction entre mendélisme pur et hypothèse chromosomique. Par contraste, l'analyse que je propose des pratiques représentationnelles caractéristiques du mendélisme et de l'étude des chromosomes contribue à souligner cette distinction, et permet d'analyser la manière dont les différentes hypothèses constitutives de la génétique classique s'articulent.

Dans la section 1, consacrée à la présentation des composantes théoriques de la génétique classique, je propose d'isoler, au sein de la génétique classique, ce que j'appellerai le mendélisme « pur », qui n'inclut pas l'explication des phénomènes génétiques au niveau cytologique, en montrant qu'il est identifiable à une méthode expérimentale et à un symbolisme introduits par Mendel et raffinés par les généticiens du début du vingtième siècle. Par contraste, je montrerai que l'explication cytologique des phénomènes génétiques repose sur des méthodes et l'utilisation de représentations d'un type différent.

Je cherche ensuite (section 2) à montrer qu'une telle approche de l'activité théorique permet d'aborder un problème classique comme celui des relations inter-théoriques de façon nouvelle. Pour cela, je présente la manière dont Philip Kitcher (1984) propose de caractériser la génétique classique et ce qu'il en appelle les « versions ». Bien que mon approche soit, sous plusieurs aspects, l'héritière de la sienne, je montrerai que sa démarche reste empreinte d'une conception des théories dont la mienne se démarque plus nettement. Je suggérerai en particulier que l'analyse des types de représentations utilisées par les théoriciens, qui conduit à distinguer nettement entre le mendélisme pur et la théorie chromosomique, permet d'aborder la question de la diversité des versions de la génétique classique d'une manière différente de celle de Kitcher. L'étude de cas du chapitre suivant mettra ainsi en évidence, par l'analyse des débats relatifs à la construction et à l'interprétation des cartes génétiques dans

les années 1920, une grande diversité d'approches du mendélisme et de l'hypothèse chromosomique au sein même de la communauté des généticiens, à une même époque.

1 Les composantes théoriques de la génétique classique

Le champ disciplinaire de la génétique classique est né en 1900, avec la redécouverte simultanée par trois botanistes européens, Hugo de Vries (1900) en Hollande, Carl Correns (1900) en Allemagne et Erich von Tschermak (1900) en Autriche, des lois que Gregor Mendel (1866) avait établies à partir de ses expériences d'hybridation trente-quatre ans plus tôt sans que son travail soit reconnu.² Les années qui suivent sont celles de l'institutionnalisation de la discipline³, de la clarification des hypothèses théoriques et du renforcement des fondements méthodologiques du nouveau programme de recherche. William Bateson en Angleterre (voir Olby, 1987; Dunn, 1965, chapitre 6) et Lucien Cuénot en France étendent les lois de l'hérédité découvertes à propos des plantes (les expériences conduisant à dégager ces régularités sont, au départ, le fait d'hybrideurs travaillant à l'amélioration des plantes) au monde animal. C'est en 1905 que William Bateson baptise le nouveau champ « génétique » (Dunn, 1965, p. 69).⁴ Le même Bateson souligne l'importance des travaux de Mendel en traduisant en 1901 son mémoire en anglais.

La période de développement la plus intense de la génétique mendélienne, par l'école de Morgan à Columbia, également appelée « groupe de la drosophile » en raison de son choix comme organisme modèle de *Drosophila melanogaster*, la mouche du vinaigre, s'étend sur les décennies 1910, 1920 et 1930. Le développement, par ce groupe, de la théorie chromosomique de l'hérédité, selon laquelle le support matériel des gènes est constitué par les chromosomes, contribue à promouvoir l'émergence de la génétique classique telle qu'on l'enseigne encore aujourd'hui.⁵

Dans ce qui suit, je n'inclus pas la théorie chromosomique de l'hérédité dans ce que l'on pourrait appeler le mendélisme « pur ». Je réserverai l'expression « génétique mendélienne » pour désigner ce dernier, et je parlerai de « génétique classique » pour l'analyse classique générale des phénomènes héréditaires, incluant leur explication au

²Pour une étude des raisons pour lesquelles le travail de Mendel a été négligé par ses contemporains, voir par exemple (Brannigan, 1985; Macroberts, 1985; Weinstein, 1962). En ce qui concerne le « mendélisme » de Mendel, voir (Olby, 1985).

³Pour une analyse du contexte social qui permet la naissance de la génétique comme discipline, voir (Gayon, 1995).

⁴Ce n'est que quatre ans plus tard que le terme « gène » est introduit par Willem Joannsen (1909) pour désigner l'unité hypothétique de transmission des caractères.

⁵Pour le détail des changements théoriques introduits de 1900 à 1920 par les généticiens classiques, voir (Darden, 1991).

niveau cytologique, par opposition à l'analyse moléculaire.⁶ On verra en effet que les composantes théoriques principales de ce que Morgan (1928) appelle « la théorie du gène » sont conceptuellement indépendantes de l'hypothèse chromosomique. L'articulation de ces deux ensembles d'hypothèses théoriques dans la pratique des généticiens des années 1920 est un des objets principaux de ce chapitre et du suivant.

Avant de les présenter, je propose d'examiner brièvement l'apport qui a effectivement été celui de Gregor Mendel. Mon propos n'est pas d'évaluer dans quelle mesure Mendel était mendélien ni d'analyser les raisons de la négligence de son travail par ses contemporains. Cependant, une des hypothèses fondamentales des arguments de ce chapitre et du suivant est que Mendel, quelle que fût sa théorie de l'hérédité, a introduit une méthode expérimentale et surtout un symbolisme destiné à en présenter et en expliquer les résultats, qui sont caractéristiques de la génétique mendélienne tout au long de son histoire. Prêter attention à cette méthode et à ce symbolisme permet d'isoler ce que j'appelle le « mendélisme pur ».

Après avoir défini la méthode et le symbolisme mendéliens (section 1.1), je présenterai (section 1.2) les composantes théoriques de la théorie du gène, c'est-à-dire de ce que j'appelle le mendélisme « pur », puis (section 1.3), la théorie chromosomique de l'hérédité. Enfin (section 1.4), j'examinerai le type de représentation utilisé en cytologie et je le comparerai avec le symbolisme mendélien. Cela me permettra d'identifier les types de représentations propres aux deux composantes théoriques principales de la génétique classique, le mendélisme « pur » et la théorie chromosomique de l'hérédité.

1.1 Méthode expérimentale et symbolisme mendéliens

Mendel procède à des expériences de croisement ou d'hybridation sur des plantes, en particulier le petit pois *Pisum sativum*.⁷ De l'observation des caractères apparents des individus des générations successives issues de ces croisements, Mendel établit les proportions des fréquences relatives de différents caractères. Pour le croisement de deux espèces⁸ « souches » présentant une différence constante concernant un caractère

⁶Ce choix est uniquement destiné à la clarté de mon propos ; il ne prétend pas suivre l'usage courant qui, à ma connaissance, consiste à employer l'une et l'autre expressions indifféremment.

⁷Ce choix est guidé par un remarquable talent d'expérimentateur : sur ce point, voir le mémoire original de Mendel (1866) et (Giordan, 1989).

⁸Il convient de souligner que le terme d'espèce n'a pas, ici, tout à fait le même sens qu'aujourd'hui (et il n'est pas même stabilisé aujourd'hui). Ce dont il s'agit dans les expériences de Mendel s'appellerait aujourd'hui « souches consanguines », « lignées consanguines » ou encore « lignées pures », c'est-à-dire le produit d'une succession de croisements consanguins destinés à réduire la variabilité génétique à zéro (ou presque) et permettre une grande stabilité des caractères au travers des

(par exemple, la couleur du pois, qui peut être verte ou jaune)⁹, il constate, à la première génération d'hybrides F_1 , que tous les individus présentent le même état de caractère (jaune). À la deuxième génération d'hybrides F_2 (issue de l'autofécondation de F_1), il observe une proportion de trois individus de couleur jaune pour un individu de couleur verte.

L'autofécondation, à son tour, de la génération F_2 permet à Mendel de constater la « réapparition » de l'état de caractère vert dans les générations issues d'individus présentant l'état de caractère jaune. Cette observation conduit Mendel à supposer la transmission d'éléments pouvant rester à l'état latent dans les cellules germinales et se manifester aux générations ultérieures. Son innovation essentielle, qui a valu aux lois de l'hérédité leur nom de « mendéliennes », est l'analyse qu'il propose alors du ratio de 3 pour 1 observé à la génération F_2 . Supposant que, parmi les individus présentant la couleur jaune, certains portent un élément à l'état latent responsable de la réapparition de l'état de caractère vert dans leur descendance, il décompose le ratio 3 : 1 en 1 : 2 : 1. Cela signifie que sur les trois individus jaunes, un seul porte, dans ses cellules germinales, le caractère souche, et les deux autres, tout en ayant la même apparence (en présentant l'état de caractère jaune), portent en fait les éléments responsables des deux états de caractère dans leurs cellules germinales.¹⁰ Ainsi, Mendel propose par là-même d'analyser les données numériques obtenues par expériences de croisement comme significatives d'un niveau d'organisation non observable.

En outre, pour expliquer ce ratio qu'il analyse comme cachant un autre ratio, Mendel introduit, en 1865, le symbolisme qui va rester celui de la génétique mendélienne. Il propose de représenter les différents types de cellules germinales par des lettres, en fonction des caractères (ou des éléments responsables des caractères) qu'elles portent. Pour un même caractère, une même lettre, en majuscule ou minuscule pour les deux états de caractères différents, par exemple, A et a . Il appelle « dominants » les caractères qui « sont transmis entièrement, de manière presque inchangée au cours de l'hybridation, et qui constituent par conséquent les caractères de l'hybride » et « récessifs » « ceux qui deviennent latents » (Mendel, 1866).¹¹ Chez

croisements entre individus d'une même lignée. Les espèces (quelle que soit leur définition), sont aujourd'hui perçues comme des entités naturelles (et non des créations artificielles) au sein desquelles la variabilité génétique est possible. Merci à Patrice David pour ces précisions.

⁹Pour plus de clarté, on distingue habituellement le « caractère » (par exemple la couleur ou la forme du pois) de l'« état de caractère » (par exemple, vert ou jaune, lisse ou ridé). Cette terminologie n'est cependant pas celle de Mendel, et il est difficile de la maintenir rigoureusement pour décrire ses travaux, comme il apparaîtra bientôt.

¹⁰Comme on va le voir, il est difficile d'établir clairement la manière dont Mendel concevait cet élément latent.

¹¹« L'expression "récessif" a été choisie parce que les caractères ainsi désignés se retirent ou dispa-

les hybrides qui présentent le caractère dominant mais chez qui le caractère récessif existe à l'état latent, il propose de désigner la cellule germinale par « Aa ». Il explique donc le ratio de 3 pour 1 par la formule suivante : $A + 2Aa + a$.¹²

Mendel se contente de ce symbolisme mathématique élémentaire, et on ne trouve aucune figure dans son mémoire. Plus tard, les diagrammes appelés « carrés de Punnett » permettront une compréhension plus évidente de ce résultat en le présentant sous la forme d'un tableau à double entrée (voir figure 36). Cette présentation sous forme de tableau implique cependant l'adoption d'une hypothèse supplémentaire, celle de l'existence de « facteurs » présents par paires dans chaque cellule.

Mendel ne fait pas explicitement une telle hypothèse. La question de savoir s'il concevait l'existence de tels facteurs, c'est-à-dire d'éléments sous-jacents dans les cellules germinales, est débattue. Les lettres qu'il emploie réfèrent explicitement à des types de cellules germinales et il est difficile de savoir comment Mendel se représentait l'existence du caractère latent dans ces cellules. Olby (1985) soutient que Mendel s'occupait des caractères observables et n'a pas proposé d'explication qui postule des paires de facteurs mutuellement exclusifs pour chaque caractère. D'ailleurs, il convient de noter que, pour les cellules « souches » (« homozygotes » dans le vocabulaire de la génétique), il emploie une seule et unique lettre ; c'est Bateson qui, en 1902, introduit la double lettre majuscule ou minuscule pour désigner les homozygotes dominants ou récessifs, sans mentionner que Mendel ne le faisait pas. Quelles que soient les conceptions qui étaient effectivement celles de Mendel, notons que l'introduction du symbolisme $A + 2Aa + a$ pour exprimer les ratio observés *permet* une interprétation en termes d'éléments présents dans les cellules germinales ; ce symbolisme est au cœur du développement ultérieur de la génétique mendélienne.

Par la suite, les acteurs de la génétique classique se livreront à des réflexions et des débats explicites sur la meilleure manière d'adapter ce symbolisme aux différentes hypothèses théoriques et aux découvertes empiriques. Par exemple, William Bateson, proposant l'hypothèse de la « présence-absence », selon laquelle l'état de caractère récessif correspond à l'absence du facteur responsable de l'état de caractère dominant propose de remplacer l'expression « Aa » par l'expression « $A0$ ». Pour rendre compte des cas d'allélomorphes multiples, c'est-à-dire des cas où il existe plus de deux états

raissent entièrement dans les hybrides, mais réapparaissent néanmoins de manière inchangée dans leur descendance » (Mendel, 1866).

¹²Une des difficultés de lecture des textes de Mendel est que le symbolisme ne reflète pas de manière évidente la distinction entre génotype et phénotype, c'est-à-dire entre les gènes responsables des caractères (les « éléments latents » chez Mendel) et les caractères apparents eux-mêmes. Aujourd'hui, on note entre crochets le phénotype – par exemple $[A]$ – et entre parenthèses ou sans rien le génotype – par exemple (Aa) . Merci à Patrice David pour cette remarque.

possibles pour un caractère, l'école de Morgan remplacera l'expression « Aa » par une lettre représentant le caractère, accompagnée d'un « + » quand il s'agit du type sauvage (Morgan 1913).¹³

Ces différents choix ne remettent cependant pas en cause, en général, l'idée même qui est au cœur de ce symbolisme, celle d'une combinatoire des particules héréditaires, particules que l'on peut faire intervenir dans les calculs sans que cela implique d'hypothèse définie sur leur nature. Quels que soient les symboles que l'on choisit¹⁴, et quelles que soient les éventuelles hypothèses sur la nature de leurs référents (les gènes), l'utilisation des mathématiques combinatoires dans l'analyse du mode de transmission des caractères héréditaires n'est pas remise en cause. Comme on va le voir, un des aspects essentiels de la conception mendélienne de l'hérédité réside dans cette combinatoire des particules héréditaires, mise en pratique dans les expériences de croisement reposant sur le choix de caractères différentiels et dans le symbolisme permettant d'en exposer les résultats.

Je propose donc, comme hypothèse de travail, de considérer le formalisme minimal introduit par Mendel et développé par ses redécouvreurs comme exprimant le cœur de la génétique mendélienne. Je l'appellerai « symbolisme mendélien » et je désignerai par « technique » ou « méthode expérimentale mendélienne » les expériences d'hybridation basées sur la sélection de paires de caractères différentiels.

1.2 La théorie du gène dans les années 1920

Au cours des dernières décennies du dix-neuvième siècle, on assiste au développement de quelques unes des conceptions fondamentales qui structurent la biologie actuelle et conduisent, au tournant du siècle, à la redécouverte des lois de Mendel.¹⁵ En particulier, le concept d'hérédité prend progressivement le sens qu'il a dans le cadre de la génétique mendélienne, lui permettant de devenir un champ de recherche autonome.¹⁶ C'est à celle-ci, telle que Morgan (1928) la présente au terme de deux décennies de recherche, sous le nom de « théorie du gène », que je vais m'intéresser

¹³Voir les réflexions de Sturtevant (1965, p. 53) ainsi que de Castle (1913) qui s'interroge sur la constance des gènes et la nécessité d'employer des lettres pour les désigner.

¹⁴Pour les besoins de l'enseignement et de la vulgarisation, il est courant que les lettres soient remplacées par des icônes, représentant par exemple des pois de couleur jaune ou verte. Cela ne modifie pas la signification des hypothèses ainsi exprimées.

¹⁵Pour une histoire de la découverte des lois de l'hérédité, voir (Lenay, 1990) qui reproduit des textes fondateurs de Charles Naudin, Gregor Mendel, Charles Darwin, August Weismann, Hugo de Vries et Lucien Cuénot en langue française, assortis d'un riche appareil critique.

¹⁶Pour une histoire du concept d'hérédité et l'analyse de manière dont il prend son sens actuel au dix-neuvième siècle, voir (Gayon, 1994, 1995, 1997, 2000).

ici. Dans ce qui suit, j'en expose les composantes théoriques fondamentales, en leur apportant un rapide éclairage historique.¹⁷

1.2.1 La distinction entre génotype et phénotype

Une des composantes fondamentales de la théorie classique de l'hérédité est la distinction entre la constitution génétique d'un organisme et l'ensemble de ses caractères observables, jointe à l'hypothèse selon laquelle ceux-ci sont déterminés par celle-là. La distinction entre des éléments « latents » transmis par les parents, dont l'existence « ne nous est connue que par ses effets sur la postériorité de l'individu » et les éléments « patents » ou apparents est introduite dans les années 1870 par Francis Galton (1872, 1875), contre l'hypothèse de la pangenèse¹⁸ de son cousin Darwin (1875). C'est une dizaine d'années plus tard que le zoologiste allemand August Weismann, aujourd'hui célèbre pour son rejet de l'hypothèse de l'hérédité des caractères acquis, propose une autre terminologie en opposant le *soma*, c'est-à-dire les caractères apparents du corps développé, et le *germen*, les entités génétiques héréditaires transmissibles par les cellules germinales, qui sont la cause des caractères apparents.¹⁹ Finalement, Wilhelm

¹⁷Je ne détaillerai pas les aspects historiques et je n'en garde que ce qui est indispensable à la compréhension de mon propos. Pour des comptes rendus détaillés de l'histoire de la génétique classique, je renvoie aux sources qui ont été les miennes. Pour la période de la redécouverte et une approche historique conceptuelle de l'ensemble du champ, voir (Gayon, 1994, 1995, 1997, 2000; Giordan, 1989; Lenay, 1990). Pour la génétique classique, voir la littérature historique secondaire (Darden, 1991; Carlson, 1967; Kohler, 1994; Olby, 1985; Mayr, 1982), ainsi que les histoires de type « *whiggish* » écrites par les principaux acteurs (Dunn, 1965; Muller, 1951; Sturtevant, 1965; Allen, 1978); en ce qui concerne la littérature première, je citerai mes sources au fur et à mesure.

¹⁸Selon cette hypothèse, chacune des cellules d'un organisme émet en permanence des « gemmules » qui circulent dans tout le corps. Certaines d'entre elles se retrouvent dans les organes reproducteurs, tout en conservant les propriétés distinctives de la cellule dont elles sont issues, et gardent la capacité d'engendrer des cellules de ce type. Lors de la fécondation, les gamètes mélangent leurs gemmules; la genèse de l'embryon consiste ensuite dans le développement des potentialités des gemmules. Dans la mesure où les gemmules ne sont pas des particules héréditaires, mais de petits « bourgeons » produits à chaque instant par chacune des cellules, et où toutes les variations, selon Darwin, sont héritées, l'hypothèse de la pangenèse implique l'hérédité des caractères acquis : « lorsqu'un nouveau caractère apparaît, quelle que soit sa nature, il tend généralement à être hérité, au moins temporairement, et parfois de manière plus persistante. » (Darwin, 1875, vol. 1, chap. 12) Je reviendrai sur un autre sens important de cette hypothèse à la section suivante. Voir (Gayon, 1994, 1995, 1997).

¹⁹Weismann voit une séparation nette entre deux substances : le « plasma germinal » (ou « germinatif »), porteur des caractères héréditaires et le « plasma somatique », où les caractères se réalisent concrètement dans la causalité actuelle de l'environnement. La continuité du plasma germinal entre les générations explique la transmission des caractères. Les variations qui se produisent dans le plasma somatique n'ont pas d'effet sur le plasma germinal; elles ne peuvent donc pas être héredi-

Joannsen (1909), en introduisant les termes « phénotype » et « génotype »²⁰, fait de cette distinction un des principes fondamentaux de la génétique.²¹

Cette « doctrine exacte, expérimentale de l'hérédité » (Joannsen, 1909, p. 1) consiste en effet à inférer le génotype des individus étudiés à partir de l'observation de leur phénotype et de celui de leur descendance, obtenue au moyen d'expériences de croisement (ou hybridation). De ce point de vue, la relation de causalité entre génotype et phénotype est fondamentale. Cependant, la manière dont le génotype, joint à l'environnement, cause effectivement le phénotype – le mode d'action des gènes – ne fait pas partie du domaine d'enquête mendélien. C'est là un deuxième aspect essentiel du mendélisme, qui se concentre sur la transmission des facteurs héréditaires (les gènes) et non pas sur leur fonction ni sur le développement de l'organisme dans son environnement. La distinction entre génotype et phénotype s'accompagne donc, dans la méthodologie mendélienne, de la séparation de l'étude de l'hérédité et de la transmission des gènes d'une part de celle du développement et du mécanisme d'action des gènes d'autre part.

Michel Morange (2003) suggère que cette séparation était indispensable, d'un double point de vue institutionnel et méthodologique, au développement de la génétique²², le mode d'action des gènes n'étant véritablement expliqué que dans le cadre de la génétique moléculaire. Il montre que les travaux ultérieurs et les développements encore très récents de la compréhension du rôle des gènes dans le développement justifient *a posteriori* l'attitude prudente de Morgan qui était « persuadé, dans les années trente, qu'étant donné la complexité des liens unissant les gènes aux

taires. Les caractères acquis ne sont pas transmissibles, puisque le corps (*soma*) est le résultat du développement des cellules germinales mais n'a pas d'influence, en retour, sur le plasma germinal. Cette distinction opérée par Weismann avait déjà été suggérée par Pierre-Philippe de Vilморin et son fils Louis de Vilморin, puis par Galton (voir Gayon, 1995).

²⁰À propos de cette distinction, voir (Lewontin, 1992).

²¹Comme me l'a signalé Patrice David, il convient cependant de ne pas assimiler la distinction génotype-phénotype avec la distinction plasma germinal-plasma somatique. En effet, le plasma germinal, pour Weismann, est une matière aussi bien que le plasma somatique; c'est la continuité matérielle du plasma germinal au travers des cellules germinales (c'est-à-dire sexuelles), et son étanchéité matérielle avec le soma, qui expliquent l'impossibilité de transmettre des caractères acquis par l'influence de l'environnement sur le soma. Mais il s'agit bien de matière et non d'information abstraite (comme l'est le génotype). On peut très bien actuellement, grâce au clonage, parler du génotype d'une cellule somatique, ou, dans les cas de recombinaisons mitotiques chez la drosophile par exemple, parler d'une cellule somatique de génotype *XX* voisinant avec une autre de génotype *YYY*.

²²Seuls des « rebelles », critiques de la génétique formelle, comme Charles Waddington et Richard Goldschmidt, s'intéressent au rôle des gènes dans le développement embryonnaire (Morange, 2003, p. 35).

caractères, l'étude de ces liens était encore prématurée » (Morange, 2003, p. 36).²³

La distinction entre les problèmes de l'hérédité et ceux du développement est en effet explicitement assumée par Morgan comme une nécessité méthodologique, qu'il présente comme la meilleure voie d'accès à une compréhension ultérieure du mécanisme d'action des gènes :

À plusieurs occasions, j'ai insisté sur l'importance de distinguer, *au moins pour le moment*, les questions relatives à la distribution des gènes dans les générations successives de celles relatives à l'action physiologique des facteurs génétiques au cours du développement [...]. On a pu affirmer qu'une telle procédure limite le champ légitime de l'hérédité à un processus qui ne va pas plus loin que celui d'un jeu de cartes, puisque le mendélisme se réduit à mélanger le jeu à et repartir avec une nouvelle « main » à chaque génération. Mon plaidoyer est, je le crains, fondé en grande partie sur des considérations d'opportunité, qu'il serait trop facile d'interpréter comme une étroitesse d'esprit ; pour autant, j'espère être parmi les premiers à accueillir une contribution réelle concernant la nature des gènes, fondée sur les changements chimiques qui se produisent dans l'embryon où les produits des gènes montrent leurs effets. En fait, je ne connais pas de voie plus directe qui puisse nous faire espérer, à terme, des découvertes concernant la nature des entités matérielles que nous concevons comme des gènes dans les cellules germinales. (Morgan, 1917)

Comme on va le voir maintenant, cette séparation de l'étude de l'hérédité et de celle du développement s'accompagne, dans les travaux des mendéliens, d'une conception des gènes qui en fait des unités abstraites, et qui se manifeste, elle aussi, dans le symbolisme utilisé pour représenter les phénomènes héréditaires.

1.2.2 Les gènes : entités discrètes et stables

Un autre aspect fondamental du mendélisme réside en effet, comme je l'ai évoqué dans ma présentation du symbolisme introduit par Mendel, dans la manière dont on représente les entités génétiques héréditaires, supposées causer les caractères apparents. Ce symbolisme acquiert la signification qu'il a dans la génétique mendélienne proprement dite au terme d'une évolution conceptuelle mise en lumière par Jean Gayon (1997, 1994, 1995, 2000), qui marque le passage d'une conception de l'hérédité comme *force* à une conception de l'hérédité comme *structure*. Cette dernière conception s'accompagne, dans le mendélisme, d'une représentation des éléments héréditaires du « *germen* » sous la forme de particules discrètes et stables, dont la structure de l'assemblage détermine les caractères apparents des organismes.

Selon Jean Gayon (1994, 1997, 2000), un des principaux promoteurs de cette

²³À ce sujet, Michel Morange renvoie à (Allen, 1978, chap. 9).

conception est Charles Darwin, critique vigoureux de l'idée d'hérédité comme « force vive ». Son hypothèse « provisionnelle » de la pangenèse, selon laquelle chacune des cellules d'un organisme vivant génère, tout au long de sa vie, des « gemmules » qui en retiennent les caractères distinctifs et se retrouvent en partie dans les cellules germinales, jouant alors le rôle de particules héréditaires, contient en réalité deux hypothèses distinctes.

La première est que les gemmules sont produites par les cellules somatiques, d'où elles sont transportées vers les cellules germinales. De ce point de vue, l'hypothèse de la pangenèse « est la forme la plus extrême d'hérédité des caractères acquis qui ait jamais été formulée » (Gayon, 1994).²⁴

La seconde est une « hypothèse de nature particulière sur le mode de transmission des caractères » (Gayon, 1994), selon laquelle « les cellules germinales contiennent un matériau spécial qui enferme les caractères héréditaires de l'organisme » (Gayon, 1997). Cette seconde hypothèse, jointe aux progrès de la cytologie, est, selon Jean Gayon (1997) à l'origine du développement de la notion de plasma germinal par Weismann (1889) et de Vries (1889) et de leurs spéculations sur les particules héréditaires. La notion de « pangènes », introduite par de Vries (1889) pour désigner les déterminants héréditaires contenus dans les noyaux des cellules germinales, « supposés être de petits organites doués des propriétés fondamentales de la vie » (Gayon, 1997), est à l'origine du terme de « gène » apparu en 1909. Ainsi, une fois évacuée la thèse de l'hérédité des caractères acquis par l'établissement d'une distinction nette entre plasma germinal et cellules somatiques, l'hypothèse darwinienne de la pangenèse rend possible la conception de déterminants héréditaires discrets et stables.

Débarassée de son aspect lamarckien (les « pangènes » devenant de pures unités de transmission [...]), et réinterprétée dans un langage mendélien, l'hypothèse « provisionnelle » de Darwin a constitué la matrice ontologique (et terminologique) des « gènes ». (Gayon, 1994)

Comme le souligne Jean Gayon (1994, 1997, 2000), décrire le plasma germinal comme constitué de déterminants particuliers permet d'aborder le problème de l'hérédité comme celui de la structure selon laquelle ces déterminants sont assemblés à chaque génération, plutôt que comme celui de leur passé. Cette conception de l'hérédité comme structure d'éléments particuliers s'oppose à une conception de l'hérédité comme force d'intensité variable, dont l'étude implique la prise en compte du passé de cette force : selon cette conception, défendue et mathématisée au moyen de la statistique par les biométriciens Francis Galton et Karl Pearson (1898), plus un caractère a été transmis sans hybridation, plus la force héréditaire – dont les biométriciens font

²⁴Voir note 18.

une grandeur mesurable – est puissante.²⁵

[...] une fois intégrées dans les cellules germinales, les gemmules darwiniennes y ont le statut d'unités de transmission, échantillonnées et recombinaées à chaque génération. Dès lors, ce qui importe, c'est la structure, et non le passé de ces assemblages. (Gayon, 1994)

Pour reprendre l'expression de Morgan (1917) citée précédemment (page 436), le mendélisme conçoit bien, en un sens, l'hérédité comme un problème consistant à « repartir avec une nouvelle “main” à chaque génération ».

Quelle que fût, encore une fois, la signification que Mendel attribuait effectivement aux symboles qu'il utilisait, sa méthode expérimentale et son recours aux mathématiques combinatoires pour étudier les phénomènes héréditaires offraient bien le cadre adéquat pour développer et exprimer une théorie consistant à « considérer le matériel génétique non pas comme une grandeur mesurable, mais comme une structure spécifique, qui s'exprime dans des croisements » (Gayon, 2000, pp. 77-79).

[La méthodologie mendélienne] a mis au premier plan l'idée d'une combinatoire des caractères, en lieu et place d'une spéculation sur le lignage. Dans la vision prémendélienne de l'hérédité, celle-ci est quasiment synonyme de « descendance » (ou, en français moderne, « l'ascendance », le réseau des ancêtres). Le concept d'hérédité ancestrale [...] revient à penser l'hérédité d'un individu comme sommation des influences reçues de tous les ancêtres. Le parti-pris méthodologique du mendélisme est totalement différent. Un mendélien s'intéresse avant tout à la structure génotypique des individus. Une fois établie pour deux parents, cette structure suffit à déterminer en nature et en proportions les structures génotypiques des enfants. (Gayon, 1994)

En outre, comme je l'ai signalé, l'emploi de ce symbolisme n'implique aucune hypothèse définie sur la nature des déterminants héréditaires : si l'objet du mendélisme est bien la transmission de ces déterminants constituant le génotype des individus, ils ne peuvent pas être observés et restent, dans le cadre de la génétique classique, des entités hypothétiques dont la nature matérielle et la structure sont inconnues. Ainsi définie, la génétique mendélienne a un caractère éminemment « abstrait », au sens où les entités dont elle parle et dont elle décrit le comportement afin d'expliquer la transmission des caractères apparents des individus sont elles-mêmes des entités abstraites. Les prédictions et les explications génétiques reposent sur la manipulation de symboles qui font référence à des entités abstraites et inobservables ; les différentes hypothèses possibles sur la nature de ces entités n'ont pas de rôle à jouer dans les explications génétiques.

Jusqu'à la révolution moléculaire, le gène est donc un « être de raison, [...] qui se

²⁵Sur le programme des biométriciens, voir (Gayon, 1992, 1996a, 1997).

présente comme une entité sans corps, sans épaisseur, sans substance » (Jacob, 1970, p. 246), dont les effets sont étudiés au moyen des résultats visibles et quantifiables des expériences de croisement. Des aspects identifiables du phénotype sont supposés déterminés par les gènes dont le mécanisme d'action est une boîte noire, et sont utilisés comme des indicateurs de la structure du génotype.²⁶

Comme on le verra, la théorie chromosomique de l'hérédité propose une explication mécanique de la transmission de ces facteurs, en affirmant que les chromosomes en sont le support matériel ; cependant, elle ne dit encore rien de la nature chimique du gène et de son mécanisme de l'action. Comme l'affirme Morgan lors de son discours de réception du prix Nobel en 1933, le niveau d'analyse de la théorie mendélienne ne requiert même pas que le gène ait effectivement une existence matérielle : « Au niveau auquel les expériences génétiques se situent, que le gène soit une unité hypothétique ou une particule matérielle ne fait pas la moindre différence. » (Morgan, 1935, pp. 347-348)²⁷

Pour résumer ce qui précède, la méthode mendélienne repose sur le traitement quantitatif²⁸ de données obtenues par des expériences de croisement (ou hybridation). Ces données permettent d'inférer le génotype d'une génération d'individus à partir des données relatives aux phénotypes de leur descendance. Les facteurs héréditaires supposés responsables des caractères apparents sont représentés comme des entités discrètes et stables, sans aucune hypothèse sur leur nature et leur mode d'action. Le mode de transmission de ces facteurs de génération en génération est décrit au moyen de deux principes, appelés « lois de Mendel ».²⁹

²⁶ « Dans cette pratique, des aspects identifiables du phénotype, supposés directement déterminés par les gènes selon un mécanisme consciemment laissé dans une boîte noire, étaient utilisés comme des indicateurs ou des "fenêtres" permettant d'avoir un aperçu de la structure formelle du génotype. » (Rheinberger et Müller-Wille, 2007)

²⁷ Les recherches des années 1930 aux années 1950 se caractérisent par une grande incertitude sur la nature et le mode d'action du matériel génétique et par plusieurs découvertes qui ébranlent certaines des conceptions classiques du gène et appellent l'explication au niveau moléculaire. Ces modifications ne remettent pas fondamentalement en cause, cependant, les principes mendéliens de l'hérédité, tels qu'on les enseigne encore aujourd'hui. Pour une histoire du concept de gène, dans le cadre classique, puis dans le cadre moléculaire, voir (Morange, 2003 ; Rheinberger, 2000 ; Kitcher, 1982a).

²⁸ Par là, je veux dire que les données relatives aux phénotypes des individus sont traitées quantitativement. Cette expression n'est pas à comprendre comme désignant la génétique dite « quantitative », qui est la génétique des caractères dont *l'observation* passe par une mesure.

²⁹ Ces deux lois ne figurent pas en tant que telles dans le mémoire de Mendel. Le premier principe est présenté comme une « loi » ou encore « règle » à partir des mémoires des redécouvreurs en 1900. Le second ne sera explicitement distingué du premier que dans les années 1910.

1.2.3 Les « lois de Mendel »

Chez les organismes pluricellulaires à reproduction sexuée, chaque individu possède un stock de gènes en double exemplaire, résultat d'une participation égale des deux parents. Les différents exemplaires d'un même gène ont été baptisés « alléomorphes » par Bateson, puis plus simplement « allèles ». Un individu possédant deux exemplaires identiques du même gène est appelé « homozygote » pour ce gène ; dans le cas contraire, il est « hétérozygote » (cette terminologie est introduite par Bateson et Saunders, 1902).

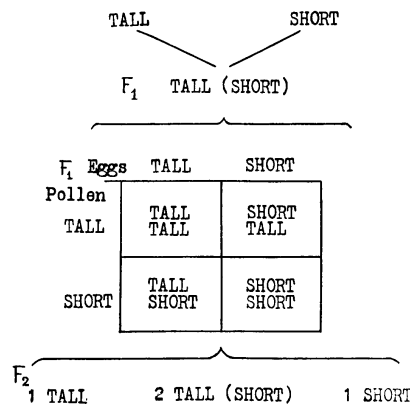


FIG. 36 – Expression diagrammatique de la première loi de Mendel par l'exemple du croisement de pois de grande taille (*tall*) avec des pois de petite taille (*short*). La génération parentale est représentée tout en haut. Juste en dessous, à la première génération d'hybrides F_1 , tous les individus sont hétérozygotes ; ils présentent donc le caractère dominant (le caractère récessif est placé entre parenthèse). Le résultat de l'autofécondation de la génération F_1 pour la recombinaison des gamètes à la génération F_2 est présenté dans le carré, expliquant la proportion de trois pois de grande taille pour un pois de petite taille à la génération F_2 (ligne du bas). Diagramme tiré de (Morgan, 1928, p. 2).

La *première loi de Mendel*, dite « loi de ségrégation » ou « de pureté des gamètes », affirme que, lors de la formation des cellules germinales (gamètes), les deux allèles de chacun des gènes sont séparés de sorte que chaque gamète ne contient qu'un allèle pour chaque gène (et a une chance égale de contenir l'un ou l'autre allèle). La progéniture reçoit donc un allèle de chacun de ses parents. Dans les cas les plus simples où un caractère est effectivement déterminé par un gène et où la dominance est claire³⁰, cette loi prédit que, lors du croisement de deux lignées pures³¹ présentant une différence constante de caractère (par exemple la taille, grande ou petite, du pois,

³⁰Je reviens, après l'exposé des deux lois, sur le statut de l'opposition entre dominance et récessivité.

³¹Voir ci-dessus, note 8.

voir figure 36), dont les individus sont par conséquent chacun homozygotes pour ce gène, toute la progéniture de la première génération d'hybride (appelée « F_1 ») présentera le caractère dominant ; elle prédit aussi que, à la génération suivante F_2 , après autofécondation des individus de la génération F_1 , on trouvera un rapport de 3 caractères dominants (grande taille) à 1 récessif (petite taille), qui s'explique par le fait que l'on trouve un homozygote dominant pour un homozygote récessif et deux hétérozygotes (présentant le caractère dominant). Si l'on désigne les deux allèles d'un même gène par la même lettre, l'une en majuscule et l'autre en minuscule, après autofécondation des hybrides, on trouve la proportion $AA + 2Aa + aa$.

		Eggs			
		GW	Gw	gW	gw
Sperm	GW	GW GW	Gw GW	gW GW	gw GW
	Gw	GW Gw	Gw Gw	gW Gw	gw Gw
	gW	GW gW	Gw gW	gW gW	gw gW
	gw	GW gw	Gw gw	gW gw	gw gw

FIG. 37 — Expression diagrammatique de la deuxième loi de Mendel, par l'exemple du croisement dihybride (couleur et forme) de *Pisum sativum*. Chacune des seize cases du diagramme représente des génotypes et phénotypes également probables de la génération F_2 . Diagramme tiré de (Morgan, 1928, p. 9).

La *deuxième loi de Mendel*, dite « loi d'assortiment indépendant », est en fait une extension de la loi de ségrégation aux cas de croisements où plusieurs caractères différentiels sont étudiés (croisements « dihybrides » pour deux caractères, par exemple, la couleur verte ou jaune du pois, et son aspect lisse ou ridé, voir figure 37). Elle affirme que les différentes paires de gènes, au moment de la formation des gamètes, s'assortissent indépendamment les unes des autres ; autrement dit, la ségrégation d'une paire de gènes n'a pas d'influence sur la ségrégation d'une autre paire. Symboliquement, pour deux paires de gènes désignées par les lettres A et B , on trouve $9AB : 3Ab : 3aB : ab$, ce qui s'explique clairement par les combinaisons mises en évidence dans le schéma de la figure 37.

L'*opposition entre dominance et récessivité*, dont on vient de voir qu'elle est assez indispensable à l'exposé des lois de Mendel et des prédictions qu'elles permettent, est parfois appelée « troisième loi de Mendel ». Comme on l'a vu, elle est introduite en 1865 par Mendel pour rendre compte du fait que des caractères « invisibles » chez

les hybrides peuvent « réapparaître » aux générations suivantes. Elle va cependant progressivement disparaître du mendélisme. L'hypothèse, chez Mendel, est que, chez les hybrides, c'est-à-dire chez des individus ne possédant pas les caractères « purs » ou « caractères souches », les deux caractères existent dans la cellule germinale, l'un s'exprimant (le caractère dominant) et l'autre restant à l'état latent (le caractère récessif), ce qui donne aux hybrides la même apparence que les individus d'une des deux espèces pures (celle possédant le caractère dominant). La discrimination entre un hybride et un individu possédant un caractère souche dominant se fait par l'examen des phénotypes de sa descendance : en effet, le croisement d'un hybride avec un individu homozygote récessif doit donner naissance à une certaine proportion d'individus présentant le caractère récessif.

Dans le cadre de la théorie du gène proprement dite, la distinction entre dominance et récessivité n'a pas vraiment sa place, pour plusieurs raisons. D'abord, on trouve de nombreux cas de dominance partielle, qui se manifestent par des caractères intermédiaires (par exemple, dans l'espèce *Antirrhinum majus*, les fleurs issues d'un croisement entre fleurs rouges et fleurs blanches sont de couleur rose). Notons que cela n'échappait pas à Mendel : dans son mémoire, il affirme avoir choisi *Pisum sativum* et les sept caractères différentiels en raison du fait que, lors du croisement, les hybrides présentaient des caractères tout à fait identiques à l'un des caractères souches, ou alors si proches que l'on ne pouvait les en distinguer, et que l'autre caractère souche restait tout à fait invisible. Ensuite, de nombreux gènes ont plus de deux formes différentes (alléomorphes multiples). Enfin, il apparaît que l'hypothèse selon laquelle chaque gène est responsable d'un caractère est une simplification très abusive de la théorie mendélienne : on trouve des gènes responsables de caractères multiples (pléiotropie) ainsi que, inversement, des caractères qui sont le résultat de l'action combinée de plusieurs gènes (polygénie).

Comme le dit Richard Lewontin (1992, p. 138), l'hypothèse de dominance – récessivité est une hypothèse sur le développement, une hypothèse épigénétique. Dans la théorie du gène, telle que Morgan (1928) l'énonce, la distinction n'est pas mentionnée (voir ci-dessous, page 444). La théorie du gène, comportant les deux lois dites de Mendel, concerne uniquement la transmission des facteurs héréditaires et non pas l'apparence finale des organismes, conséquence de plusieurs processus causaux qui ne dépendent pas des lois d'hérédité du matériel génétique.³²

³²C'est ce qui permet à Kenneth Waters (1990) d'affirmer que, contrairement à une simplification erronée habituellement admise, le dogme central de la génétique classique n'est pas que les gènes déterminent directement les caractères, mais qu'une différence génotypique cause une différence phénotypique, par des mécanismes dont la complexité ne sont pas l'objet de la génétique classique, mais deviendront ceux de l'analyse moléculaire. D'ailleurs, l'hypothèse de dominance – récessivité

1.2.4 Liaison partielle et crossing-over

La dernière composante théorique principale de la génétique mendélienne des années 1920 est tout à fait absente du mémoire de Mendel ; elle est constituée de la notion de liaison partielle et de l'hypothèse, proposée par Morgan au début des années 1910, dite du « crossing-over ». Le développement de cette hypothèse, comme on le verra au chapitre suivant, est à la fois le moteur et la conséquence de l'adoption et de la défense, par le groupe de Morgan, de la théorie chromosomique de l'hérédité. Elle est en est cependant conceptuellement indépendante.

À partir de 1905-1906, plusieurs expériences de croisement révèlent que certains gènes tendent à être hérités ensemble (c'est-à-dire qu'on les retrouve liés dans plus de 50% des cas, ce qui est la proportion attendue en raison de la loi d'assortiment indépendant), sans pour autant que cette liaison soit complète (100%) ; on parle alors de *liaison partielle*. Les gènes ainsi liés sont considérés comme faisant partie d'un « groupe de liaison ». Quand ils ne sont pas hérités ensemble, on dit qu'il y a eu une *recombinaison*. Une fois les phénomènes de liaison partielle reconnus, la deuxième loi de Mendel doit être amendée : elle ne s'applique qu'aux gènes faisant partie de groupes de liaison différents.³³

Les phénomènes de liaison partielle font l'objet, dès leur découverte, de nombreuses hypothèses différentes de la part des généticiens, sur lesquelles je reviendrai plus en détail au prochain chapitre. Je ne présente, ici, que l'explication proposée par Morgan, celle du *crossing-over*. Selon cette hypothèse, un échange ordonné se produit parfois entre les portions correspondantes de deux groupes de liaison homologues, expliquant les cas de liaison à plus de 50% et moins de 100%. En raison de l'observation de certaines régularités dans les phénomènes de liaison partielle, sur lesquelles je reviendrai en détail au prochain chapitre, Morgan affirme en outre que les gènes de chaque groupes de liaison sont ordonnés linéairement, de telle sorte que la fréquence des crossing-over entre deux gènes est un indice de la distance qui les sépare. Cette suggestion de Morgan, qui est à l'origine de l'invention de la technique

trouve une explication au niveau moléculaire, en termes de quantité d'effet de gènes et de processus de régulation. Comme me l'a signalé Patrice David, la « dominance » d'une paire d'allèles est elle-même une quantité susceptible d'évoluer et / ou de changer en fonction de l'environnement ; un allèle d'un gène *A* peut être clairement dominant sur *a* dans un certain contexte et codominant ou récessif dans un autre contexte (par exemple en présence d'un allèle nouveau au gène *B*, ou dans un environnement plus chaud ou plus froid. De même, dans le cas d'un gène à effet pléiotrope sur deux caractères, un allèle peut être dominant pour un caractère et récessif pour l'autre.

³³Lindley Darden (1991, p. 139) remarque que la deuxième loi n'est explicitement distinguée de la première qu'à partir du moment où on lui trouve des exceptions. L'expression « assortiment indépendant » apparaît pour la première fois en 1913. Morgan *et al.* (1915) ne parlent pas encore de deuxième loi ; on la trouve explicitement mentionnée dans (Morgan, 1919).

de cartographie génétique par son élève Alfred Sturtevant, est directement liée à son adoption de la théorie chromosomique de l'hérédité. Cette dernière en offre, comme on va le voir ci-dessous (section 1.3), une explication mécanique. Cependant, l'hypothèse du crossing-over peut être adoptée comme une hypothèse génétique sur la recombinaison des gènes sans que cela implique l'adoption de l'hypothèse chromosomique. Notons d'ailleurs que, dans son énoncé de la « théorie du gène », qui résume tous les éléments que je viens de présenter, Morgan (1928) ne fait aucune mention des chromosomes ; les groupes de liaison pourraient avoir une base matérielle d'une tout autre nature :

La théorie affirme que les caractères d'un individu sont dus à des paires d'éléments (gènes) dans le matériel germinal, lesquelles sont unies dans un certain nombre de groupes de liaison ; elle affirme que les membres de chaque paire de gènes se séparent quand les cellules germinales sont formées, en accord avec la première loi de Mendel, et en conséquence que chaque cellule germinale en vient à contenir un seul ensemble d'éléments ; elle affirme que les membres appartenant à différents groupes de liaison s'assortissent indépendamment en accord avec la deuxième loi de Mendel ; elle affirme qu'un échange ordonné – le crossing-over – se produit quelquefois entre les éléments de groupes de liaison correspondants ; et elle affirme que la fréquence des crossing-over fournit la preuve de l'ordonnement linéaire des éléments de chaque groupe de liaison et de la position relative des éléments relativement les uns aux autres. (Morgan, 1928, p. 25)

1.3 La théorie chromosomique de l'hérédité

Malgré l'indépendance des principes mendéliens de leur explication au niveau de la cellule, considérée à la fin du dix-neuvième siècle comme la base de l'organisation du vivant, le besoin d'expliquer les phénomènes héréditaires à partir d'un support physique s'est fait sentir dès les débuts de la génétique mendélienne. L'hypothèse, suggérée à plusieurs reprises dès la fin du dix-neuvième siècle³⁴, qui va progressivement s'imposer et donner finalement lieu à un nouveau programme de recherche dès les années 1910, est la théorie chromosomique de l'hérédité, selon laquelle le matériel héréditaire dont les lois de Mendel décrivent la transmission entre générations se trouve sur – ou est identifiable à – des parties de chromosomes. L'interprétation de cette

³⁴Lindley Darden (1991) compte au moins trois théories chromosomiques proposées au cours de l'histoire. La première est une hypothèse purement cytologique proposée à la fin du dix-neuvième siècle par Naegeli ainsi que par Wilson : selon cette hypothèse, la chromatine porte le matériel héréditaire (l'« idioplasme » selon Naegeli). La deuxième est la théorie du plasma germinal de Weismann, qui s'accompagne du rejet de l'hérédité des caractères acquis. Enfin, la troisième, celle qui va être appelée à se développer et à devenir la théorie admise de façon quasi unanime à la fin des années 1920, est proposée par Boveri et Sutton en 1902-1903.

hypothèse de « localisation » des gènes sur les chromosomes ou « d'identification » des gènes à des portions de chromosomes varie à la fois en fonction de la manière dont on conçoit les gènes (entités matérielles, « charges », etc.) et des hypothèses sur les propriétés matérielles des chromosomes, dont les observations cytologiques ne fournissent qu'une connaissance imprécise.

1.3.1 Observations cytologiques et hypothèse chromosomique

La cytologie, c'est-à-dire l'étude du comportement des cellules constituant les êtres vivants, est une science déjà bien développée à la fin du dix-neuvième siècle, et les cytologistes ont une connaissance détaillée des étapes du processus de division cellulaire normale, la mitose.³⁵ En revanche, au tournant du siècle, les étapes du processus de formation des cellules germinales – la méiose – sont encore mal connues. Mendel s'appuie cependant sur les avancées de la cytologie et conçoit la ségrégation en termes de formations de cellules germinales.

Au cours des années 1880 à 1899, les cytologistes commencent à considérer l'importance du noyau et de la chromatine, et identifient les chromosomes. C'est en 1882 que Walter Fleming, à partir de recherches sur les amphibiens, distingue dans le noyau une partie fortement colorable, la chromatine, et une partie peu colorable, l'achromatine. Il met en évidence que les granules de chromatine se disposent pendant la mitose en filaments appelés, en 1888, « chromosomes ». Walter Fleming établit également que, pendant la mitose, les chromosomes se dupliquent et se divisent longitudinalement. S'opposent alors deux théories, qu'aucune observation cytologique ne permet de départager, à propos des chromosomes. Selon la première, défendue par les cytologistes Hertwig et Fick, les chromosomes se dissolvent entièrement à la fin du processus de division cellulaire pour se reformer au début du nouveau cycle. Une seconde thèse, qui sera reprise et développée par Boveri, affirme l'individualité et l'intégrité des chromosomes tout au long de la vie de la cellule.³⁶

En 1901, Montgomery apporte un premier argument en faveur de la continuité du chromosome, repris par Sutton en 1902 : l'observation montre que certains chromosomes morphologiquement individualisables pendant la division cellulaire sont repérables avec les mêmes caractéristiques à chaque division cellulaire. Les observations de

³⁵En 1860, la constitution de la cellule, considérée comme l'unité de base de l'organisation du vivant, est présentée par Max Schulze comme « une petite masse de protoplasme avec un noyau ». En 1875, le botaniste allemand Edward Strasburger met en évidence le phénomène de la mitose en montrant que, chez les plantes, les noyaux dérivent toujours d'autres noyaux. La même année, Bütschli, puis Boveri, mettent en évidence le même phénomène de mitose chez les animaux.

³⁶Pour plus de détails techniques et bibliographiques concernant la période historique brièvement retracée dans cette section, voir (Dunn, 1965).

Montgomery et de Sutton montrent que pendant la première prophase de la méiose, deux chromosomes semblables s'accolent pour se séparer avant la réalisation de la division réductrice. Ils suggèrent alors que l'ensemble des chromosomes d'une espèce consiste en paires de chromosomes homologues, un représentant de chaque paire étant d'origine maternelle et l'autre d'origine paternelle. Sutton ne tarde pas à conclure que « l'association par paires des chromosomes paternels et maternels et leur séparation lors de la méiose pourraient constituer la base physique des lois mendéliennes de l'hérédité ».

Cependant, il convient de souligner que la théorie de Sutton, défendue également par Boveri, ainsi que par Cannon et de Vries, n'a qu'un statut hypothétique ; elle est suggérée par l'analogie entre, d'une part, la présence des chromosomes par paires, chaque exemplaire étant hérité d'un des deux parents, et, d'autre part, les lois concernant la transmission des facteurs mendéliens. Aucune observation cytologique ne permet de confirmer l'hypothèse selon laquelle les chromosomes sont effectivement le support du matériel héréditaire, et de nombreuses questions restent en suspens concernant la relation entre les différences morphologiques des chromosomes et les différences génétiques.

En particulier, la thèse de Boveri et Sutton de l'individualité des chromosomes, aux sens morphologique et fonctionnel du terme, rend inconcevable le fait que les chromosomes se brisent et échangent des parties. Elle prévoit par conséquent que les caractères dont les facteurs sont portés par le même chromosome sont hérités ensemble. Selon Sutton, en effet, l'allélomorphe est une partie du chromosome ; il ne peut pas, par conséquent, en envisager un échange. La loi d'assortiment indépendant semble donc contredire la théorie de Sutton. En revanche, Hugo de Vries suppose que les « pangènes » ne sont pas réellement des parties de chromosomes, mais sont « portés » par les chromosomes, qui peuvent échanger leurs pangènes, sans que cela consiste à renoncer à l'indivisibilité des pangènes. On a pu voir dans cette hypothèse de De Vries une première version de la théorie du crossing-over. Cependant, l'hypothèse de l'échange de pangènes n'est portée par aucune observation cytologique ; de plus, contrairement à l'hypothèse du crossing-over, elle n'est pas motivée par des phénomènes de liaison partielle, non encore constatés. Elle est plutôt un amendement à la théorie chromosomique, preuve que cette dernière a encore du mal à s'accorder aux lois mendéliennes, qui prévoient l'indépendance de la ségrégation des différents facteurs mendéliens. Nombreux sont les mendéliens qui, avant 1910, s'opposent à la théorie chromosomique. Le plus illustre des généticiens classiques, Thomas H. Morgan, en est un farouche opposant jusqu'en 1910.

Un élément essentiel dans les progrès de la cytologie et dans son rapprochement de la science de l'hérédité est l'affirmation de la *détermination chromosomique du sexe*.

Dès 1891, les observations cytologiques permettent de repérer un morceau de chromatine non pair, que Wilson appelle « X ». En 1905, Wilson et Stevens confirment l'hypothèse de la détermination chromosomique du sexe proposée en 1902 par McClung, en observant que les femelles de plusieurs insectes ont deux chromosomes X , là où les mâles semblent n'en avoir qu'un.

La période de 1906 à 1910 voit le renforcement de l'hypothèse chromosomique, mais les opinions des généticiens divergent encore fortement. William Bateson, avocat du mendélisme, considère les phénomènes chromosomiques comme un effet, au même titre que d'autres différences dans les organismes, de différences héréditaires plus profondes. Morgan (1910a) lui-même, qui va devenir le grand avocat et développeur de l'hypothèse chromosomique à partir de 1911, admet la détermination chromosomique du sexe mais refuse l'hypothèse selon laquelle les gènes sont localisés sur les chromosomes.³⁷

1.3.2 Janssens (1909) et la chiasmotypie. L'adoption de l'hypothèse chromosomique par Morgan.

C'est au début des années 1910 que Morgan adopte la théorie chromosomique. Cette « conversion » est due, en parallèle de la découverte de plus en plus nombreux cas de liaison partielle, à l'observation par le cytologiste Janssens (1909) de processus d'entrecroisements entre les chromosomes homologues pendant une des premières phases de la méiose. Janssens (1909) propose alors la théorie de la *chiasmotypie*,

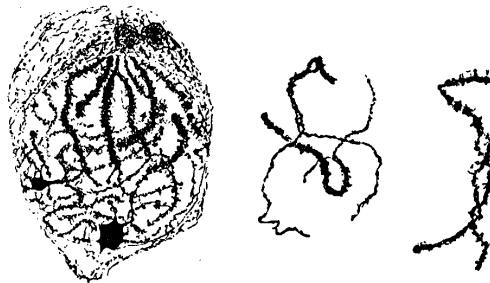


FIG. 38 – Entrecroisement de chromosomes chez l'amphibien *Batrachoseps* pendant la méiose. (Janssens, 1909)

selon laquelle, au cours de ces entrelacs, les chromosomes homologues se cassent et échangent certaines de leurs parties. Cette hypothèse n'est soutenue par aucune preuve observationnelle : la seule chose visible est l'entrecroisement des chromosomes, mais rien ne permet de conclure à un échange effectif de portions de chromosomes.

³⁷Pour l'évolution intellectuelle de Morgan, voir (Carlson, 1967; Allen, 1978).

Il faudra attendre les années 1930 pour que des observations cytologiques confirment cette hypothèse.

Conscient du caractère hypothétique de la théorie de la chiasmatypie, Morgan y voit cependant la promesse d'une explication mécanique aux phénomènes de liaison partielle ; la chiasmatypie (l'échange de portions entre chromosomes homologues) serait ainsi la base physique des phénomènes de crossing-over (l'échange de facteurs entre groupes de liaison homologues). En retour, les phénomènes de liaison partielle invitent à voir dans les entrelacs de chromosomes l'indice d'un tel échange ; en ce sens, si les observations cytologiques semblent pouvoir apporter une explication mécanique aux phénomènes génétiques, ces derniers suggèrent une certaine lecture des micrographes de chromosomes. Comme le remarque Morgan (1928) au chapitre consacré au « mécanisme de l'hérédité mendélienne », cette lecture est d'autant plus probable que les cytologistes ont suggéré la possibilité d'une chiasmatypie avant même de connaître les règles de transmission des gènes :

[...] bien que l'on doive admettre que la preuve cytologique du crossing-over n'a pas été apportée, [...] il a néanmoins été montré dans un grand nombre de cas que les chromosomes se trouvent dans une position dans laquelle un tel échange peut être raisonnablement supposé se produire. [...]

Le cytologiste [...] nous présente les chromosomes d'une manière qui remplit de nombreuses exigences de la génétique. Quand on rappelle que nombre de ces preuves ont été obtenues avant la redécouverte du mémoire de Mendel, et que ce travail n'a jamais été accompli avec un biais génétique, mais bien indépendamment de ce que les chercheurs étudiant l'hérédité faisaient, il ne semble pas probable que ces relations soient de pures coïncidences, mais plutôt que les chercheurs étudiant les cellules ont découvert plusieurs aspects essentiels du mécanisme par lequel les éléments héréditaires sont triés suivant les deux lois de Mendel, et sont échangés de manière ordonnée entre les membres de la même paire de chromosomes. (Morgan, 1928, p. 44)

Sur ces affirmations prudentes, Morgan et son groupe bâtissent une explication mécanique – qui reste hypothétique jusqu'à sa confirmation dans les années 1930 – des phénomènes génétiques de crossing-over, qu'ils présentent au moyen d'un schéma, baptisé « modèle du collier de perles » (figure 39). Ce modèle représente la structure du chromosome comme celle d'un assemblage de « perles », supposées représenter les gènes. Malgré son caractère analogique – le groupe de Morgan ne prétend rien affirmer sur la nature matérielle de ces « perles » – ce modèle comporte une hypothèse importante : celle de l'*arrangement linéaire des gènes le long du chromosome*, qui signifie que chaque gène est « attaché » à chacun de ses deux voisins, mais à aucun autre, la totalité du chromosome se présentant donc comme une chaîne. En représentant ainsi l'arrangement des gènes le long du chromosome, le groupe de Morgan suggère

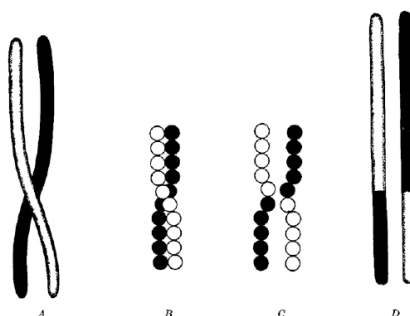


FIG. 39 – Modèle du « collier de perles » expliquant les phénomènes de crossing-over. Schéma tiré de (Morgan *et al.*, 1915, p. 60).

que les phénomènes de crossing-over doivent pouvoir être décrits comme des phénomènes mécaniques : cela permet de comprendre plusieurs propriétés importantes des phénomènes de liaison partielle, que je présenterai au chapitre suivant.³⁸

Les travaux de Morgan et de ses élèves dans les années 1910 et 1920 sont portés par l'adoption de ce modèle ; de nombreuses découvertes – génétiques et cytologiques – semblent offrir des indices de plus en plus convaincants en faveur de l'hypothèse chromosomique. Cependant, jusque dans les années 1930, aucune observation cytologique ne donne accès à la structure fine des chromosomes ni ne met en évidence des échanges effectifs de portions au cours de la méiose. Jusqu'à cette période, l'explication cytologique des phénomènes génétiques reste hypothétique ; le modèle des perles sur un fil n'est donc lui-même qu'une analogie heuristique, que le groupe de Morgan se défend, à plusieurs reprises, d'interpréter de manière réaliste. Je reviendrai sur cette période de manière plus détaillée au prochain chapitre, où j'analyserai la manière dont les engagements de différents généticiens des années 1920 en faveur de l'hypothèse chromosomique – ou de son rejet – jouent dans leur compréhension même du mendélisme.

³⁸En tout premier lieu, l'hypothèse de linéarité permet de comprendre que, dans certains cas, si l'on considère trois gènes A , B et C , la fréquence de recombinaison entre deux gènes A et C se trouve être la somme des fréquences de recombinaison entre A et B et entre B et C . On considère alors que ces trois gènes se trouvent sur la même ligne, B étant situé entre A et C . On verra au chapitre suivant que cette additivité souffre des exceptions, qui seront elles-mêmes expliquées mécaniquement au moyen du modèle du collier de perles.

1.4 Types de représentation en cytologie et en génétique mendélienne

J'ai suggéré, à la section 1.1, que le symbolisme et la méthode combinatoire introduits par Mendel permettent de caractériser l'activité théorique propre au mendélisme, que j'ai distingué de la génétique classique, en tant que cette dernière englobe l'hypothèse chromosomique. Cela signifie que les explications mendéliennes sont typiquement obtenues et exprimées au moyen de ce type de symbolisme, qu'on le présente linguistiquement sous la forme d'équations ou au moyen de tableaux comme les carrés de Punnett, qui sont des extensions bidimensionnelles du symbolisme mendélien permettant de calculer les chances de répartition des phénotypes des enfants étant donné le génotype des parents. De même, les concepts de la génétique mendélienne, comme ceux de liaison partielle et de crossing-over, peuvent être exprimés au moyen de ce symbolisme : la liaison partielle est caractérisée par une certaine fréquence d'association de deux gènes, et un crossing-over est identifié à la recombinaison de facteurs habituellement liés.

Afin de bien distinguer le niveau d'explication – ou de description – proprement mendélien du niveau d'explication cytologique, je propose, dans cette section, de m'arrêter un instant sur le type de représentations typiquement utilisées en cytologie, ainsi que sur la manière dont elles sont convoquées par les généticiens pour expliquer les phénomènes génétiques. Cette comparaison jouera un rôle important dans l'argument de la section 2 du présent chapitre ; l'analyse de la relation entre ces types de représentation dans la pratique des généticiens des années 1920 sera au cœur de mon analyse de la cartographie génétique dans le chapitre suivant.

La cytologie, « la science qui cherche à explorer l'espace de la cellule » (Jacob, 1970, p. 229), repose, à la période qui m'intéresse, sur l'utilisation du microscope optique et l'emploi de certains composés qui colorent sélectivement différents tissus de la cellule, permettant d'en étudier la structure ainsi que de suivre les étapes de processus dynamiques comme la mitose et la méiose.

Contrairement aux représentations utilisées en génétique mendélienne, les images obtenues au microscope sont des représentations d'entités *concrètes* (cellules, chromosomes, etc.) ; les processus que le microscope permet de visualiser sont des processus *spatio-temporels*. Pour reprendre la terminologie définie au chapitre 7 (section 1.3.2), ces représentations sont *picturales* : les relations spatiales au sein de la représentation tiennent lieu de relations spatiales au sein de l'objet étudié. En l'absence d'un accès à la structure fine de la cellule – que la microscopie électronique offrira plus tard – on peut établir que les chromosomes, identifiés grâce à des techniques de coloration, prennent, à un certain stade de la division cellulaire, la forme de « bâtonnets » qui

se dupliquent et se divisent (voir figure 40).

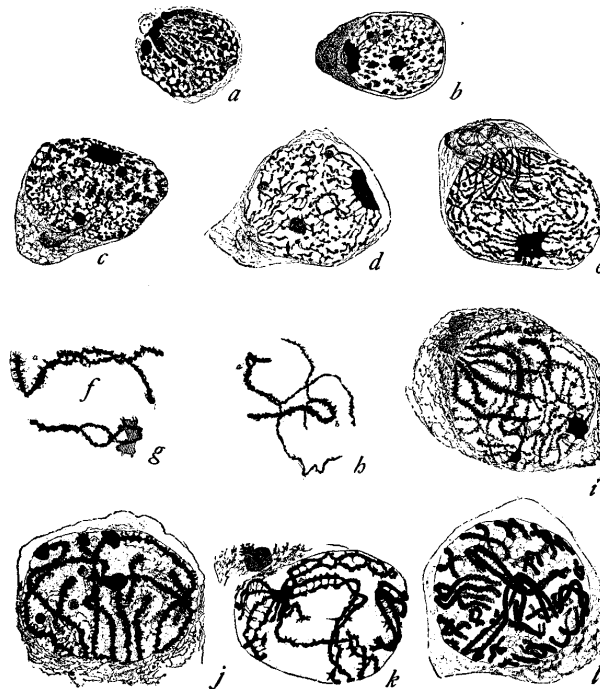


FIG. 40 — Spermatogénèse chez le batracien. *a* : télophase des dernières divisions spermatogoniales; *b* : étape de repos après la dernière division spermatogonale; *c* : apparition de filaments minces donnant au noyau un aspect spirématique; *d* et *e* : étape du bouquet grêle (le filament spirématique devient évident et s'oriente vers un côté du noyau); *f*, *g*, *h* : étape du bouquet amphitène (détail : formation de filaments épais); *i* : étape du bouquet amphitène (cellule entière); *j* : étape du bouquet pachytène; *k* : étape strepsinème (séparation longitudinale des fils); *l* : raccourcissement et épaississement des chromosomes. (Janssens, 1909, image et légende tirées de)

Ces représentations picturales, en tant qu'elles sont obtenues par des techniques élaborées de coloration des différents tissus, contiennent déjà plusieurs hypothèses sur la composition de la cellule. On peut ensuite en tirer des représentations plus ou moins schématiques, permettant de mettre en évidence ces hypothèses (c'est déjà le cas des images de la figure 40, et en particulier des images *f*, *g* et *h* qui présentent les chromosomes isolés de la cellule). On accentuera ainsi la « ligne » qui sépare l'intérieur de l'extérieur de la cellule en lui donnant éventuellement la forme d'un cercle, on augmentera le contraste entre différents éléments en gommant des traces jugées peu significatives, et on ajoutera des symboles linguistiques indiquant la nature des différents éléments. Ainsi, par exemple, dans la figure 41, les cellules sont représentées par des cercles, au sein desquels un autre cercle figure le noyau; le cytoplasme est grisé et le noyau conservé blanc, afin de rendre les chromosomes plus visibles. Ces der-

niers sont d'ailleurs représentés au nombre de trois, et on distingue les chromosomes paternels des chromosomes maternels par leur coloration. Comme je l'ai suggéré au chapitre 7 (section 1.3.3), si les représentations picturales sont souvent les sources de ces représentations schématiques, ces dernières apprennent en retour à lire les images picturales en fonction des hypothèses qu'elles expriment. Quelles que soient

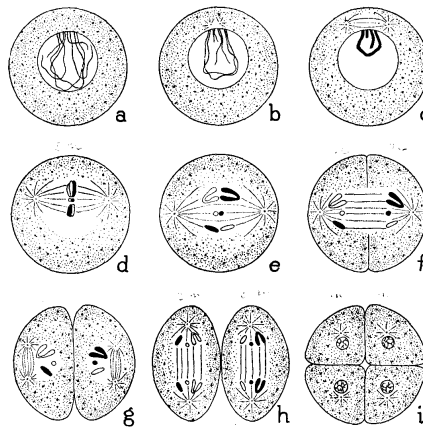


FIG. 41 – Schéma représentant deux étapes de la méiose. Seules trois paires de chromosomes sont représentées ; les chromosomes paternels sont colorés en noir, les chromosomes maternels en blanc (sauf dans les figures *a*, *b* et *c*). L'étape de la méiose réductionnelle est représentée dans les figures *d*, *e* et *f*. La méiose équationnelle est représentée dans les figures *g*, *h* et *i*. (Morgan, 1928, p. 33)

les abstractions et les distorsions opérées lors des schématisations progressives à partir des images obtenues au microscope, ces représentations conservent la propriété importante de respecter les relations topologiques de la chose représentée ; on peut les projeter sur des images picturales. En ce sens, elles restent des représentations, même générales, d'entités spatiales concrètes.

On peut, au moins partiellement, caractériser le travail théorique du cytologiste comme reposant sur la manipulation de telles représentations : par le perfectionnement des techniques d'observation et le travail d'abstraction et de généralisation à partir des micrographes, qui se concrétise dans la construction de schémas, le cytologiste élabore et développe des hypothèses théoriques sur la cellule ; au cours de ce travail, « le paysage qui se dévoile au fond du microscope prend du relief » (Jacob, 1970, p. 230).³⁹

Ainsi, les types de représentations utilisées en cytologie se distinguent nettement

³⁹Pour une analyse détaillée du travail d'abstraction et de généralisation en cytologie à partir d'images picturales à la période qui m'intéresse ici, voir l'article de Jane Maienschein (1991) sur l'utilisation des schémas par le cytologiste Wilson. Elle y montre ainsi que, dans ses premiers travaux, Wilson s'appuie principalement sur des images picturales, qu'il considère « plus proches de

de celles qui caractérisent l'activité théorique propre au mendélisme, tel que j'ai proposé de le caractériser. Il existe cependant, de part et d'autre, des « cas limites », difficiles à classer dans l'une ou l'autre catégorie. Je voudrais en présenter brièvement deux exemples, afin de suggérer la manière dont l'analyse des représentations utilisées dans la pratique est une méthode fructueuse pour aborder certaines questions relatives aux relations inter-théoriques.

Mon premier exemple est tiré des travaux de Wilson (1925) sur la cellule. Certains schémas de processus cellulaires présentés dans son ouvrage de 1925 comportent des hypothèses théoriques qui appartiennent en propre au domaine de la génétique. Ainsi, par exemple, dans la figure 42, Wilson représente le cycle de la méiose de manière hautement schématique, et y ajoute des indications d'ordre génétique (on notera en effet, sous les schémas de cellule, des symboles représentant les facteurs mendéliens). En outre, les sept schémas de la deuxième ligne, représentant les recombinaisons, se rapprochent fortement, pris ensemble, d'un diagramme mendélien, puisque chaque image représente une probabilité de recombinaison ; Wilson précise, dans la légende, que seules deux possibilités de combinaison des chromosomes sur seize, et seules sept possibilités de recombinaison sur 256 sont représentées, et renvoie à la figure 43.

Cette seconde figure, représentant la ségrégation et l'assortiment indépendant, est explicitement présentée par Wilson comme susceptible d'être « *appliquée aux chromosomes ou aux facteurs héréditaires* ». Elle peut donc être considérée comme une représentation très *schématique* des processus de division cellulaire, obtenue par abstraction à partir de représentations picturales des phases successives de la méiose ; on peut également y voir une représentation *diagrammatique* de probabilité de distribution de génotypes typiquement mendélienne (pour la distinction entre schémas et diagrammes, voir chapitre 7, sections 1.3.3 et 1.3.4).

Le jeu entre ces deux types de représentation illustre bien la rencontre de deux domaines théoriques différents : on peut y « voir » à la fois une étape de la matérialisation du gène, et l'introduction en cytologie d'hypothèses impossibles à tirer de la seule observation des cellules. La « présence » des facteurs mendéliens au sein des cellules n'est en effet pas observable ; ils ne sont d'ailleurs pas figurés par des représentations de type schématique, mais par l'introduction, au sein des schémas de cellules, de lettres appartenant typiquement au symbolisme mendélien.

Mon second exemple est celui du modèle du collier de perles, présenté précédemment (figure 39). C'est en quelque sorte le cas inverse : alors que les figures 42

la réalité » ; plus tard, devenant plus « confiant dans ses propres interprétations », il a recours à des représentations de plus en plus abstraites, qui sont elles-mêmes influencées par les travaux de Morgan sur l'hérédité. Voir aussi l'article de Lynch (1988), mentionné au chapitre 7 (section 1.3.3), sur la relation entre les images picturales et les schémas.

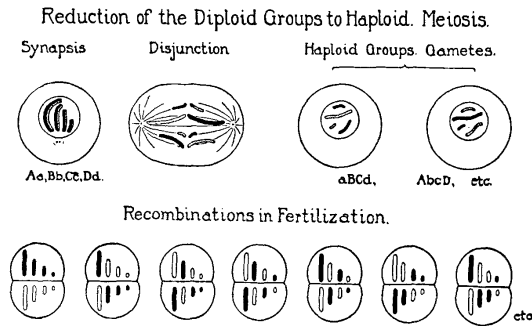


FIG. 42 – Représentation schématique de la phase réductionnelle de la méiose et de la recombinaison. Les chromosomes paternels sont représentés en blanc et les chromosomes maternels en noir. (Wilson, 1925)

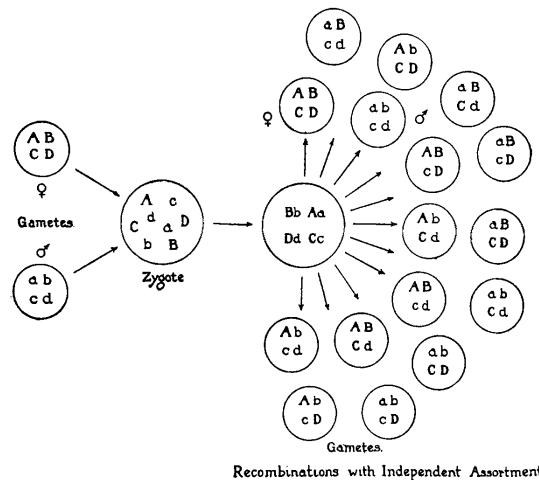


FIG. 43 – « Diagramme applicable aux chromosomes ou aux facteurs héréditaires, montrant la ségrégation et l'assortiment indépendant mendéliens. (Les phénomènes de liaison ne sont pas pris en compte ici.) Les éléments maternels (chromosomes ou facteurs) sont représentés en lettres capitales, les éléments paternels en minuscules. On suppose que le nombre de chromosomes est de quatre. » (Wilson, 1925, légende originale)

et 43 sont celles d'un cytologiste qui introduit des hypothèses génétiques dans sa représentation des cellules, ce schéma est l'œuvre de généticiens cherchant dans les connaissances en cytologie une explication de phénomènes génétiques. En outre, ce schéma est censé expliquer ce que la figure de Wilson (43) ne montre pas : les phénomènes de liaison partielle.

Le schéma de Morgan n'est pas à proprement parler tiré d'une image picturale. Il n'est pas le fruit d'une abstraction progressive qui consisterait à grossir certains traits et à en gommer d'autres. Certes, la forme générale du chromosome est le résultat

d'une telle abstraction. Mais les « perles » représentant les gènes ne sont en aucun cas le fruit d'une observation ; c'est l'introduction d'une hypothèse qui, de plus, ne doit pas être interprétée de façon réaliste. En effet, Morgan ne prétend pas que la structure du chromosome se présente effectivement sous la forme de perles. En outre, la manière dont ces perles sont attachées au chromosome ne fait l'objet d'aucune hypothèse définie : sont-elles attachées à un fil, ou bien sont-elles des parties de ce fil, entièrement constitué de perles attachées les unes aux autres ? L'analogie que suggère ce modèle entre le comportement des gènes et celui des chromosomes n'est donc pas du même ordre que celle mise en évidence par le schéma de la figure 43 : cette dernière est une analogie entre deux processus effectivement observés, au niveau cytologique (le comportement des chromosomes pendant la méiose) et au niveau génétique⁴⁰ (les règles de transmission des facteurs héréditaires, déduites de l'observation des caractères). Ici, l'analogie est seulement hypothétique : *si* l'on se figure le chromosome comme un collier de perles, on peut comprendre les phénomènes de crossing-over comme des phénomènes mécaniques.

En tant que ce modèle ne prétend rien affirmer de la structure réelle des chromosomes, il n'appartient pas au domaine de la cytologie, mais bien de la génétique. Pour autant, il n'appartient pas au type de représentations que j'ai défini comme « mendélien » (voir page 433) : il s'agit d'une représentation *schématique*, et sa valeur explicative repose sur cet aspect schématique, et non sur l'utilisation des mathématiques combinatoires. En effet, en tant qu'explication mécanique des phénomènes génétiques, il ne fonctionne que si l'on suppose qu'il représente spatialement une entité spatiale. En ce sens, l'explication mécanique du crossing-over, dans la mesure où elle repose sur un type de représentation schématique, est déjà une excursion hors du mendélisme « pur ».

Je reviendrai sur ce point dans mon analyse des débats autour de la cartographie génétique au prochain chapitre. Je tenais cependant à souligner dès à présent la manière dont différents types de représentation sont attachés à différents domaines théoriques et niveaux d'explication, car cela joue un rôle important dans l'argument de la prochaine section.

2 Identité et versions de la génétique classique selon Philip Kitcher (1984)

Ma présentation des composantes de la génétique classique suggère la chose suivante : bien plus encore que dans un exemple paradigmatique de théorie scientifique

⁴⁰ Je reviendrai, au prochain chapitre, sur le sens de la notion d'observation en génétique.

comme celui de la mécanique classique, il semble impossible de caractériser le contenu de la « théorie » mendélienne de l'hérédité au moyen d'un ensemble de principes permettant d'obtenir des prédictions et des explications des phénomènes de transmission des caractères de génération en génération.

À supposer que l'on accepte, contre ce que j'ai montré dans les chapitres précédents, de définir le contenu de la mécanique au moyen des axiomes newtoniens du mouvement, on ne peut pas raisonnablement affirmer que les « lois de Mendel » en sont les analogues dans le domaine de la génétique. Non seulement ces lois ne s'appliquent pas universellement – le phénomène de liaison partielle en est une exception notable, parmi d'autres – mais encore de nombreuses hypothèses additionnelles, liées aux découvertes empiriques des généticiens tout au long de la première moitié du vingtième siècle, interviennent dans la prédiction et l'explication de la transmission des caractères héréditaires chez les différentes espèces qu'ils étudient. S'il existe quelque chose comme une « théorie » mendélienne, on en rend mieux compte en prêtant attention aux méthodes expérimentales et aux types de représentations utilisés par les généticiens pour exprimer et développer leurs hypothèses, ainsi que pour prédire et expliquer les données relatives aux phénotypes des individus obtenus par hybridation.

Philip Kitcher (1984) dresse un constat similaire : contre la conception positiviste des théories comme des ensembles d'énoncés, il propose de caractériser la génétique classique en se référant à ce qu'il appelle la « pratique » de cette science à différentes étapes de son développement, une des composantes essentielles de cette pratique étant la mise en œuvre d'un certain « schème de raisonnement » [*pattern of reasoning*]. En outre, il s'appuie sur cette caractérisation de la génétique classique pour montrer qu'elle n'est pas réductible à la génétique moléculaire.

Dans cette section, après avoir brièvement rappelé le contexte du débat sur la réductibilité de la génétique classique à la génétique moléculaire, je présente les deux aspects de l'argumentation de Kitcher qui intéressent mon propos (section 2.1) : sa caractérisation de la génétique classique et plus particulièrement des « schèmes de raisonnement » qui lui sont propres, et l'argument principal qu'il en tire contre la thèse réductionniste. La critique que je propose de ces arguments, dans un deuxième temps (section 2.2), est double. Elle consiste d'abord à montrer que Kitcher assimile à tort les « schèmes de raisonnement » qui caractérisent selon lui l'activité théorique à un certain type d'explication, faisant par là une concession importante à l'approche positiviste des théories scientifiques qu'il prétend pourtant critiquer. Par contraste, mon analyse de la génétique mendélienne se situe dans une perspective qui se démarque nettement de celle qui aborde l'étude des relations inter-théoriques sous l'angle de la notion de réduction. Ma critique se poursuit par la mise en évidence d'un défaut dans la caractérisation même que Kitcher propose de la « pratique » de la génétique

classique.

2.1 Le débat sur la réduction et la position de Philip Kitcher (1984)

La caractérisation de la génétique classique proposée par Kitcher (1984) s'inscrit dans le cadre du débat sur la réductibilité de la génétique classique à la génétique moléculaire. Son argumentation repose sur un rejet de la conception des théories scientifiques issue du positivisme, dans les termes de laquelle la question du réductionnisme est généralement posée. Il propose donc une autre conception des théories, qu'il applique au cas de la génétique classique pour finalement montrer que, si la génétique moléculaire a bien enrichi notre compréhension de la génétique classique, celle-ci n'est pas entièrement réductible à celle-là. Je ne prétends pas prendre part, dans ce qui suit, au débat sur le réductionnisme ; les propos de Kitcher m'intéressent en tout premier lieu parce qu'ils suggèrent une approche des théories scientifiques dont la mienne est, à plusieurs égards, l'héritière. Une des critiques que je lui fais est précisément de ne pas être allé assez loin dans cette direction et de concéder trop à la conception classique des théories qui sous-tend les débats sur la réduction. Auparavant, il est utile de rappeler brièvement le contexte du débat sur le réductionnisme en génétique.

2.1.1 Le contexte du débat sur la réduction de la génétique classique à la génétique moléculaire

Le modèle positiviste de la réduction, formellement dépendant de la conception des théories qui les définit comme des ensembles de propositions, est une tentative pour formaliser l'idée selon laquelle les progrès de la science se font dans le sens d'une unification grandissante des explications fournies par les théories et tendent à une unité épistémologique, sinon ontologique. Selon le modèle classique de la réduction inter-théorique dans sa version la plus aboutie, celle d'Ernest Nagel (1961) – mentionné à plusieurs reprises dans ce travail – réduire une théorie T_2 à une théorie T_1 , c'est déduire les énoncés de T_2 des énoncés des T_1 . Pour cela, deux conditions formelles doivent être satisfaites. La première est que les *concepts* de la théorie réduite T_2 soient définissables à l'aide des concepts de T_1 , par l'ajout éventuel d'hypothèses additionnelles.⁴¹ La deuxième condition est que les *lois* de T_2 soient déductibles des

⁴¹Dans le cas paradigmatique étudié par Nagel, celui de la réduction des lois de la thermodynamique à la théorie cinétique des gaz, l'hypothèse additionnelle selon laquelle la température d'un corps est l'énergie cinétique moyenne de ses molécules permet de relier le concept thermodynamique de température aux concepts de la mécanique statistique.

lois de T_1 et par conséquent que les *phénomènes expliqués* par T_2 puissent aussi être expliqués par T_1 .

En 1969, Kenneth Schaffner (1969b) propose de décrire les avancées de la génétique sur une forme amendée de ce modèle et affirme que la génétique classique est réduite à la génétique moléculaire alors en plein essor. La découverte, par Crick et Watson en 1953, de la structure en double hélice de l'ADN, a en effet ouvert la voie d'une compréhension de la nature chimique des gènes et de leur mécanisme d'action. De nombreux phénomènes décrits, mais non expliqués, par la génétique classique, trouvent ainsi une explication au niveau moléculaire, comme par exemple la capacité des gènes à s'auto-répliquer, ainsi que la manière dont ils causent des différences phénotypiques.⁴²

Contre cette thèse réductionniste se forme au début des années 1970 chez les philosophes des sciences ce que Kenneth Waters (1990) appelle un « consensus anti-réductionniste ».⁴³ Selon Kenneth Waters (2007), deux sortes d'arguments sont avancés par ces philosophes. La première ligne argumentative consiste à montrer que les concepts de la génétique classique, et principalement le concept de gène, ne peuvent pas être définis au niveau moléculaire (la première condition formelle énoncée ci-dessus n'est donc pas satisfaite). La deuxième ligne argumentative consiste à affirmer que la dérivation des lois de la génétique classique à partir de celles de la biologie moléculaire, à *supposer qu'elle soit possible*, n'y apporterait pas d'éclairage *explicatif* ; au contraire, entrer dans les moindres détails [*gory details*] des phénomènes moléculaires ferait perdre, selon les anti-réductionnistes, la puissance explicative propre au niveau cytologique. Selon Kenneth Waters, cette ligne argumentative, à laquelle il

⁴²Chaque molécule d'ADN consiste en une double chaîne de nucléotides de quatre types différents (guanine, cytosine, thymine et adénine). Les deux chaînes de nucléotides sont enroulées l'une autour de l'autre et forment ainsi une double hélice. Des liens hydrogène relient les nucléotides adjacents deux à deux, de telle sorte que la guanine se trouve toujours en face de la cytosine et la thymine face à l'adénine. Les deux chaînes sont ainsi complémentaires. Un gène est défini (j'ignore ici toutes les difficultés liées à une telle définition) comme un segment sur l'une des chaînes. Ce modèle permet de comprendre le mécanisme par lequel le gène se réplique : quand les deux chaînes se séparent, chacune d'entre elles peut générer une nouvelle chaîne par l'association de nucléotides complémentaires. Il permet également de comprendre le mécanisme d'action des gènes : chaque chaîne sert à la synthèse de molécules d'ARN selon un mécanisme analogue à celui de la réplication ; les molécules d'ARN servent elles-mêmes à la synthèse des protéines. Pour une analyse de la manière dont la biologie moléculaire permet d'expliquer les différentes fonctions du gène, voir (Waters, 1994, 2007). Pour le récit de la découverte de la double hélice, voir (Watson, 1968). Pour une histoire de la biologie moléculaire, voir (Jacob, 1970; Morange, 2003).

⁴³Ses principaux défenseurs sont David Hull (1972, 1974, 1979), William Wimsatt (1976), Lindley Darden et Nancy Maull (1977), Philip Kitcher (1984), Alexander Rosenberg (1985, 1994), Richard Burian (1985a) et John Dupré (1993).

s'oppose, est particulièrement puissante, dans la mesure où elle est indépendante de la conception positiviste des théories.

Dans son article en faveur de la thèse anti-réductionniste, Philip Kitcher (1984) avance les deux lignes argumentatives dégagées par Waters. C'est à la seconde qu'il accorde le traitement le plus important, et c'est la seule à laquelle je m'intéresserai ici.⁴⁴

2.1.2 La pratique de la génétique classique et ses versions selon Kitcher

Un des buts de Kitcher (1984) est de proposer une conception des théories scientifiques qui permette d'analyser les relations entre ce qu'il appelle les « versions » successives d'une même théorie ainsi qu'entre théories différentes. Comme cela apparaîtra dans l'exposé que je vais faire de sa position, la notion de version telle que l'emploie Kitcher ne correspond pas à celle que j'ai définie précédemment dans ce travail (notamment au chapitre 3). Mon but est justement de montrer que la caractérisation de la génétique et de ses versions par Kitcher reste tributaire d'une conception abstraite de l'activité théorique.

Par la définition de ce qu'il appelle la « pratique » de la génétique classique à différentes époques, Kitcher cherche à fournir les outils nécessaires à l'identification de la structure de cette théorie, qu'il considère sous-jacente à ses versions successives. Identifier ainsi la structure de la génétique classique lui permet ensuite d'évaluer l'apport explicatif de la génétique moléculaire. Dans cette section, je présente les grandes lignes de sa caractérisation de la pratique de la génétique classique.

⁴⁴Kitcher montre que, même amendée, la thèse réductionniste consiste en trois exigences dont aucune n'est satisfaite par le cas de la génétique. Ces exigences sont les suivantes : « [R1] La génétique classique contient des lois générales sur la transmission des gènes qui peuvent servir de conclusions à une dérivation réductive. » « [R2] Le vocabulaire distinctif de la génétique classique (des prédicats tels que "1 est un gène", "1 est dominant par rapport à 2") peut être reliés au vocabulaire de la biologie moléculaire par des principes ponts. » « [R3] Une dérivation de principes généraux sur la transmission des gènes à partir des principes de la biologie moléculaire expliquerait pourquoi les lois de la transmission des gènes sont valides (dans la mesure où elles le sont). » (Kitcher, 1984, p. 339) [R1] n'est pas satisfaite, puisque, comme je l'ai mentionné, il n'existe pas d'ensemble d'énoncés pouvant tenir lieu de « lois » de la génétique proprement dites. L'objection des « moindres détails » est destinée à montrer que [R3], à supposer même que [R1] le soit, n'est pas satisfaite. L'argumentation de Kitcher pour montrer que [R2] n'est pas satisfaite dans le cas de la génétique repose principalement sur le problème soulevé par la signification du concept de gène. Ce problème, qui demanderait à lui seul un traitement séparé, n'est pas central dans l'article de 1984, surtout destiné à définir une approche des théories qui se démarque de l'analyse positiviste, et à développer l'objection des « moindres détails », et je ne m'en occuperai pas ici.

La pratique de la génétique classique La notion de *pratique* d'une science, développée dans (Kitcher, 1983), est assez proche, comme Kitcher le reconnaît à plusieurs reprises, de celle de paradigme chez Kuhn (1962/1970).⁴⁵ Contre l'idée positiviste d'une théorie comme ensemble d'énoncés, Kitcher définit en effet la pratique d'une science comme un ensemble hétérogène, dont les énoncés admis à une certaine époque par les praticiens d'une science, que les positivistes cherchent à reconstruire formellement, ne sont qu'une composante parmi d'autres :

Le corpus d'énoncés à propos de l'hérédité des caractères acceptés à une certaine époque n'est qu'une seule des composantes de l'entité bien plus complexe que j'appellerai la *pratique* de la génétique classique à cette époque. Il y a un langage commun utilisé pour parler des phénomènes héréditaires, un ensemble d'énoncés acceptés dans ce langage (le corpus d'énoncés sur l'hérédité mentionné ci-dessus), un ensemble de questions considérées comme les questions qu'il convient de poser à propos des phénomènes héréditaires et un ensemble de schèmes de raisonnement [*patterns of reasoning*] qui sont instantiés dans la réponse à certaines des questions appropriées ; (et aussi un ensemble de procédures expérimentales et de règles méthodologiques [...]). La pratique de la génétique classique à une époque est pleinement spécifiée si l'on identifie chacune des composantes ci-dessus. (Kitcher, 1984, p. 352)

Laissant les procédures expérimentales et les règles méthodologiques de côté, c'est sur les schèmes de raisonnement et les problèmes typiques qu'ils servent à résoudre que Kitcher se concentre. Un schème de raisonnement est défini⁴⁶ comme « une séquence d'énoncés schématiques [*schematic sentences*], c'est-à-dire des énoncés dans lesquels certains éléments du vocabulaire non logique ont été remplacés par des symboles non interprétés [*dummy letters*], jointe à un ensemble d'instructions pour les interpréter [*filling instructions*] » (Kitcher, 1984, p. 353). La notion de schème de raisonnement est destinée à présenter formellement l'idée kuhnienne selon laquelle l'apprentissage et la maîtrise d'un domaine scientifique consistent plus en l'apprentissage et la maîtrise d'un savoir-faire mis en œuvre dans la résolution de problèmes qu'en la connaissance de « lois » ou de « principes »⁴⁷ : « Cette notion de schème de

⁴⁵ Pour une comparaison entre la notion de pratique et celle de paradigme, voir (Kitcher, 1983, chap. 7).

⁴⁶ Kitcher (1981, 1989) développe en détail cette notion, indépendamment de celle de pratique, dans ses travaux sur l'explication scientifique.

⁴⁷ « On n'enseigne pas (et on n'a jamais enseigné) aux néophytes un petit nombre de lois théoriques fondamentales à partir desquelles des "théorèmes" génétiques doivent être déduits. On leur apprend une certaine terminologie technique, qui est utilisée pour avancer une grande quantité d'information sur des organismes spécifiques. Certaines questions sur l'hérédité chez ces organismes sont posées et reçoivent une réponse. *Ceux qui comprennent la théorie* [Je souligne] sont ceux qui savent quelles questions doivent être posées concernant les exemples encore non étudiés, et qui savent comment

raisonnement, ajoute Kitcher, est destinée à élucider [*explicate*] l'idée d'une structure commune sous-tendant un groupe de solutions de problèmes. » (*ibid.*, p. 353)

Sur cette base, Kitcher propose de caractériser la génétique classique en définissant les problèmes qu'elle cherche typiquement à résoudre :

tout au long de l'histoire de la génétique classique, la théorie est orientée en direction de la réponse à des questions sur la distribution de caractéristiques dans les générations successives d'une généalogie, et elle propose de répondre à ces questions en utilisant des probabilités de distributions de chromosomes pour calculer les probabilités des génotypes de la descendance. (Kitcher, 1984, p. 354)

Ces questions concernant les distributions de phénotypes (qu'il s'agisse de comprendre une distribution donnée ou de prédire les probabilités de la distribution qui sortira d'un croisement) sont ce que Kitcher propose d'appeler des « problèmes de pedigree ». En général, la génétique classique résout ces problèmes en faisant des hypothèses sur les gènes pertinents.

Chaque cas d'un problème de pedigree, explique Kitcher, peut être caractérisé par un ensemble de données (des énoncés décrivant la distribution des phénotypes parmi les organismes d'un pedigree, ou un diagramme transmettant la même information), un ensemble de contraintes (un ensemble d'informations d'ordre cytologique et la description de la constitution chromosomique des membres de l'espèce étudiée), et une question concernant les phénotypes. Résoudre un problème de ce type revient ensuite à avancer ce que Kitcher appelle une « hypothèse génétique » spécifiant les allèles pertinents, leur expression phénotypique et leur transmission au sein du pedigree et à en inférer, par un calcul, les réponses à propos des probabilités de distribution de phénotypes. Une hypothèse génétique fournit des informations de quatre types :

(a) La spécification du nombre de loci pertinents et le nombre d'allèles à chaque locus ; (b) Spécification des relations entre génotypes et phénotypes ; (c) Spécification des relations entre gènes et chromosomes, des faits concernant la transmission des chromosomes aux gamètes [...] et les détails de la formation des zygotes ; (d) Assignation de génotypes aux individus du pedigree. (Kitcher, 1984, p. 356)

Définir ainsi les problèmes typiques et les schèmes de raisonnement mis en œuvre par la génétique classique pour les résoudre permet, selon Kitcher, à la fois d'identifier ce qu'il appelle la « structure » commune à toutes les versions successives de la génétique classique et d'étudier les relations entre cette théorie et la génétique

appliquer le langage technique aux organismes concernés par ces exemples, et qui savent appliquer les schémas de raisonnement qui doivent être instantiés dans l'élaboration des réponses. Plus simplement, les étudiants qui réussissent saisissent des schémas de raisonnement généraux qu'ils peuvent utiliser pour résoudre de nouveaux cas. » (Kitcher, 1984, p. 354)

moléculaire. Selon Kitcher, la génétique classique se caractérise, tout au long de son histoire, par l'adoption de certaines expressions pour caractériser les phénomènes héréditaires, de questions d'une certaine forme, et d'un « style général de raisonnement ». Kitcher reste assez flou sur le lien entre le style général de raisonnement de la génétique classique et les différents schèmes mis en œuvre à chaque époque, qui en caractérisent les différentes versions ; il semble cependant qu'il veuille dire par là que ces schèmes ont une forme commune, mais que les hypothèses génétiques acceptables changent au fur et à mesure des découvertes empiriques, ce qui contribue à affiner les schèmes « de telle sorte que l'on s'accommode des exemples récalcitrants » (p. 355) aux versions précédentes.

Les diagrammes de pedigree comme solutions typiques de problèmes Une manière de comprendre ces propos assez abstraits de Kitcher sur les schèmes de raisonnement et les types de problèmes propres à la génétique classique, ainsi que sur le changement qu'implique selon lui une nouvelle version, consiste à considérer un type de représentations omniprésent dans les manuels de génétique : les diagrammes de pedigree. Affirmer que ces diagrammes incarnent les problèmes typiques et les schèmes de raisonnement de la génétique classique tels que Kitcher les définit ne me semble pas trahir ses propos.⁴⁸

Ces diagrammes ont la forme d'arbres généalogiques⁴⁹, où chaque génération successive est représentée sur une ligne, et où les relations d'engendrement et de fratrie sont mises en évidence par des lignes. Un ensemble de conventions – assez variables selon les cas – en permet la lecture. Par exemple, dans les figures 45, 46, 47 et 49, les hommes sont représentés par des carrés et les femmes par des cercles. Différentes colorations de ces carrés et de ces cercles indiquent, selon les cas, le phénotype ou le génotype des individus (la figure 44 présente les deux en parallèle).

Les diagrammes de pedigree peuvent servir à présenter des données particulières, comme dans le cas du pedigree de l'hémophilie de la famille de la reine Victoria (figure 47) ; dans d'autres cas, les données sont choisies pour présenter la règle de transmission de certains gènes dans des espèces particulières (voir par exemple figures 45 et 46). Les diagrammes peuvent aussi, comme dans la figure 44, servir à exprimer les lois de Mendel, au même titre que les carrés de Punnett présentés précédemment (figures 36 et 37), ou encore des règles générales comme celle de la

⁴⁸Il y fait d'ailleurs une référence explicite p. 355 : tout problème de pedigree inclut un ensemble de données, qui sont « des énoncés décrivant la distribution de phénotypes parmi les organismes d'un pedigree particulier, ou un diagramme présentant la même information [Je souligne]. »

⁴⁹Selon la typologie proposée au chapitre 7 (section 1.3.4), il s'agit bien de représentations diagrammatiques : les relations spatiales y représentent des relations de génération et de fratrie.

transmission d'un caractère lié au sexe (figure 48).

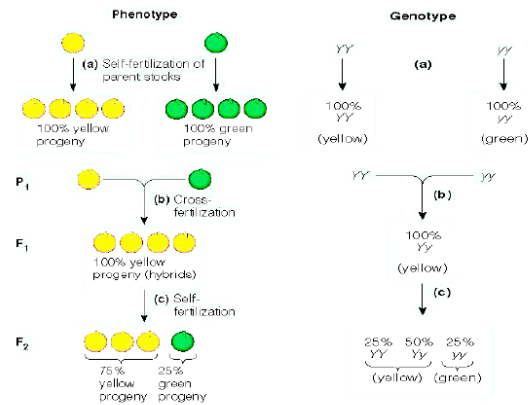


FIG. 44 – Expression de la première loi de Mendel sous la forme d'un arbre (à gauche : niveau phénotypique; à droite : niveau génotypique)

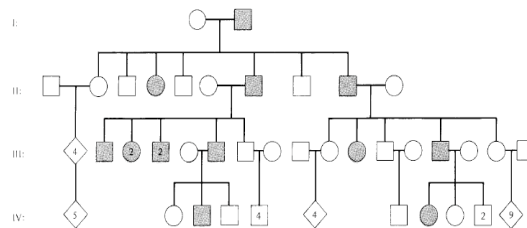


FIG. 45 – Pedigree d'un caractère dominant (nanisme chondrodystrophique) (Strickberger, 1985).

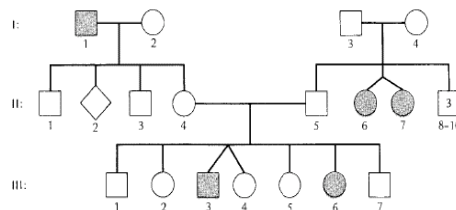


FIG. 46 – Pedigree d'un caractère récessif (albinisme) (Strickberger, 1985).

Les diagrammes peuvent donc être utilisés pour exprimer des données à différents degrés de généralité. Apprendre à raisonner en génétique, selon la caractérisation de Kitcher, consiste à apprendre quelles sont les combinaisons autorisées et les combinaisons interdites. Ainsi, l'apprentissage de la génétique commence par celui de la lecture d'un diagramme comme celui de la figure 44, qui exemplifie les règles syntaxiques élémentaires de ce type de diagramme. Ces règles, qui évoluent en fonction

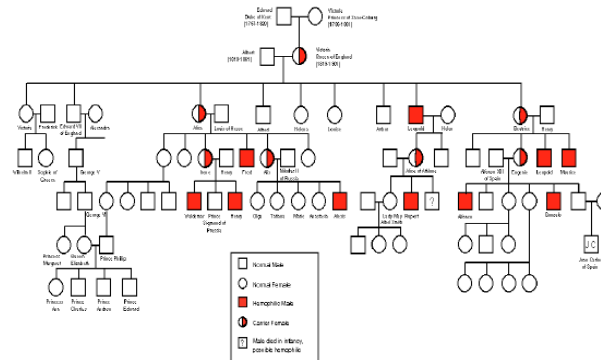


FIG. 47 — Pedigree de l'hémophilie dans la famille de la reine Victoria (<http://www.knowledgene.com>).

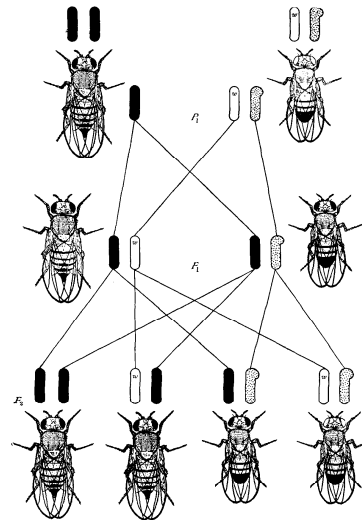


FIG. 48 — Hérité d'un caractère lié au sexe chez *Drosophila* (Morgan, 1928, p. 60).

des découvertes concernant le comportement de certains gènes et les avancées des recherches sur les chromosomes (les versions de la génétique, selon Kitcher), se complexifient au cours de l'apprentissage. Ils sont à la fois des expressions de la théorie, des représentations des données, et des « ordinateurs » permettant de tirer rapidement des inférences à partir des données d'un problème. Ainsi, après avoir appris les lois de Mendel au moyen de diagrammes comme celui de la figure 44, l'étudiant apprendra la règle de transmission des caractères liés au sexe, au moyen d'un diagramme comme celui de la figure 48.

Les règles syntaxiques de ces diagrammes sont données, dans chaque cas, par

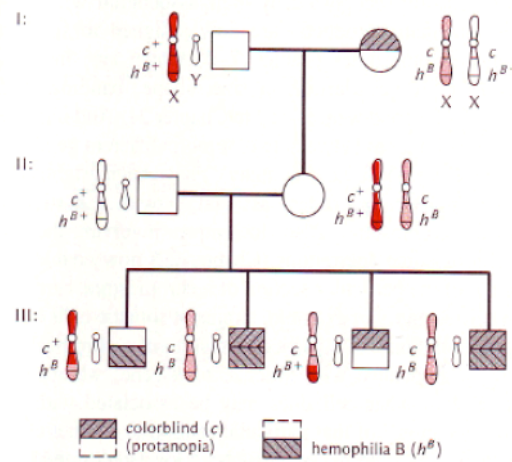


FIG. 49 – Pedigree d'une famille pour deux caractères liés au sexe, daltonisme (c) et hémophilie B (h^B). À la troisième génération, les échanges de portions de chromosomes X sont mis en évidence par la coloration des portions concernées en fonction de leur origine : rouge (grand-père $I - 1$) et rose (grand-mère $I - 2$) (Strickberger, 1985).

ce que Kitcher appelle les « contraintes » du problème, c'est-à-dire des informations relatives au mode de transmission des gènes concernés. Ces règles évoluent donc naturellement au fur et à mesure des développements de la génétique. Pour reprendre les propos de Kitcher, on peut considérer le diagramme de la figure 44 comme exprimant le « style » général de raisonnement propre à la génétique classique. La figure 48 est l'expression de la règle de transmission des caractères liés au sexe : au lieu de carrés et de ronds, les individus drosophiles sont représentés par des icônes montrant le caractère (les yeux rouges – en l'occurrence, noirs) étudié, auxquels sont ajoutées des représentations des chromosomes concernés permettant de comprendre la manière dont le caractère est transmis. Cette figure est à la fois l'expression d'une règle générale et l'exemple typique de la manière dont les problèmes de pedigree mettant en jeu un caractère lié au sexe doivent être résolus.

La maîtrise des règles syntaxiques des diagrammes à une époque donnée, et les informations relatives aux contraintes – souvent représentées au moyen d'icônes indiquant l'emplacement des gènes concernés sur les chromosomes – permettent de tirer des inférences à propos de la probabilité de distribution des gènes aux générations suivantes – ou encore à propos du génotype des ancêtres, le phénotype des descendants étant connu.

Les diagrammes permettent également de représenter certains types particuliers d'hérédité, exemplifiant ainsi le type d'inférences autorisées ou non en fonction des contraintes du problème, pour reprendre les termes de Kitcher : ainsi, la figure 48

présente les moyennes statistiques des données obtenues dans le cas de croisements mettant en jeu un caractère lié au sexe dans l'espèce *Drosophila*. Elle représente à la fois les données et la règle de transmission héréditaire concernant les caractères liés au sexe ; l'intervention d'icônes représentant des chromosomes apporte l'explication cytologique de cette règle. Elle exemplifie par là même le type de raisonnement à l'œuvre dans la résolution de problèmes de ce genre, omniprésents dans les manuels de génétique.

La figure 49, enfin, présente la transmission dans une famille particulière de deux caractères (le daltonisme et l'hémophilie), tous deux liés au sexe, mais assez éloignés sur le chromosome X. La présence d'icônes représentant des chromosomes sur lesquels on a précisé l'emplacement des gènes responsables de ces caractères permet d'expliquer, en raison de la distance séparant ces deux gènes, que la recombinaison soit de 50%. Cela permet d'expliquer comment une femme saine pour les deux caractères, fille d'un père sain et d'une mère daltonienne mais non hémophile, se reproduisant avec un homme sain et non porteur des gènes responsables du daltonisme et de l'hémophilie, peut donner naissance à quatre fils dont deux sont affectés des deux maladies et les deux autres sont l'un daltonien et l'autre hémophile. Ici, le diagramme ne représente pas la moyenne quantitative de résultats d'expériences de croisement, mais bien un cas particulier que les règles de transmissions des facteurs héréditaires permettent d'expliquer.

2.1.3 La position anti-réductionniste de Kitcher : l'explication cytologique est préférable.

Ayant ainsi posé que les problèmes typiques de la génétique classique sont des problèmes relatifs à la distribution des gènes dans les pedigrees étudiés, et que leur résolution consiste en la mise en œuvre d'un certain schème de raisonnement que j'ai proposé d'illustrer au moyen des diagrammes de pedigree, Kitcher montre que, si la génétique moléculaire a bien apporté une « extension explicative » à la génétique classique, on ne peut pas affirmer qu'elle en fournit une *réduction* explicative.

La notion d'extension explicative est destinée à rendre compte de « l'idée presque universelle selon laquelle la génétique moléculaire a fait quelque chose d'important pour la génétique classique » (Kitcher, 1984, p. 337). Examinant trois exemples d'avancées majeures opérées par la génétique moléculaire (l'explication de la répllication des gènes, la caractérisation des processus de mutation et l'identification des structures moléculaires de certains gènes d'organismes particuliers), Kitcher en conclut qu'elles contribuent le plus souvent à *expliquer des présupposés* de la génétique classique, qui, s'ils figurent à titre de prémisses implicites dans les raisonnements

typiquement mis en œuvre pour résoudre des problèmes de pedigree, ne constituent pas l'objet de cette théorie.

Disons qu'une théorie T' fournit une *extension explicative* d'une théorie T dans les cas où il existe un schème de résolution de problèmes de T dont une des prémisses schématiques peut être obtenue comme conclusion d'un schème de résolution de problème de T' . Quand une nouvelle théorie fournit une extension explicative d'une autre ancienne théorie, des prémisses particulières des dérivations explicatives de l'ancienne théorie se trouvent elles-mêmes expliquées au moyen des arguments fournis par la nouvelle théorie. Cependant, il ne s'ensuit pas que ces explications fournies par l'ancienne théorie puissent être améliorées en remplaçant les prémisses en question par les dérivations pertinentes. Ce qui est pertinent pour expliquer un énoncé S n'est pas toujours pertinent pour expliquer un énoncé S' qui figure comme une dérivation explicative de S . (Kitcher, 1984, p. 365)

Ainsi, par exemple, l'idée selon laquelle les gènes peuvent se répliquer était assez unanimement admise par les généticiens classiques, sans qu'ils soient en mesure d'expliquer ce phénomène. Son explication par la génétique moléculaire est indiscutablement, selon Kitcher, un approfondissement de la compréhension des phénomènes étudiés par la génétique classique. Cependant, pour ce qui constitue le cœur de la pratique de la génétique classique, les problèmes de pedigree, le schème de raisonnement mis en œuvre par la génétique classique – et illustré par les diagrammes de pedigree – est, selon Kitcher, irréductible au schème de raisonnement de la génétique moléculaire. Voyons rapidement en quoi consiste son argument.

Selon Kitcher, les règles de transmission des gènes de génération en génération, qui constituent l'objet de la génétique classique, ne trouvent pas d'explication satisfaisante au niveau moléculaire, alors que le niveau cytologique leur en fournit une. Prenant l'exemple simple de l'assortiment indépendant des gènes situés sur des chromosomes non homologues, il suggère qu'une explication adéquate de ce phénomène implique l'identification de ce qu'il appelle un « *processus-PS* »⁵⁰ c'est-à-dire « un processus au cours duquel des paires d'entités (en l'occurrence, des chromosomes homologues) sont séparées par une force telle que chaque membre de la paire est assigné à une entité descendante (en l'occurrence, un gamète). » Or, affirme Kitcher, si l'on peut, au niveau cytologique, identifier des processus de ce type comme formant une « espèce naturelle » (les processus de méiose), « les processus-PS ne peuvent pas être identifiés comme un même type de processus du point de vue moléculaire » (p. 349) ; en effet, les « processus-PS sont hétérogènes du point de vue moléculaire. Il n'y a pas de contrainte sur les structures moléculaires des entités qui vont par paires ou sur la

⁵⁰ « PS » signifie « *pair-separation* ».

manière dont les forces fondamentales se combinent pour les unir et les séparer. Les liens peuvent être forgés et brisés de plusieurs manières différentes » (p. 349). Bien que tous les processus décrits au niveau cytologique correspondent à des processus chimiques complexes, le niveau d'explication adéquat des règles de transmission des gènes est, selon Kitcher, le niveau cytologique. Expliquer ces phénomènes au niveau moléculaire résulte en une perte du pouvoir unificateur de l'explication par identification de processus-PS.⁵¹

La distribution des gènes dans les gamètes doit être expliquée, non pas en énumérant dans les moindres détails [*gory details*] les réarrangements [*reshuffling*] des molécules, mais par l'observation du fait que les chromosomes sont alignés par paires avant la division méiotique et que l'un des chromosomes de chaque paire est transmis à chaque gamète. On peut formuler cela dans le langage favori des biologistes en disant que l'assortiment des allèles doit être compris au niveau cytologique. Cela signifie qu'il y a un schème de raisonnement qui est appliqué pour dériver la description de l'assortiment des allèles et qui implique des prédicats qui caractérisent les cellules et leurs structures internes à une grande échelle. Ce schème de raisonnement est objectivement préférable au schème moléculaire qui serait instantié par la dérivation qui retrace les réarrangements compliqués de molécules individuelles, parce qu'il peut être appliqué dans un grand nombre de cas qui seraient hétérogènes dans la perspective moléculaire. Intuitivement, le schème cytologique établit des liens qui sont perdus au niveau moléculaire, et doit donc être préféré. (Kitcher, 1984, pp. 370-371)

Non seulement les détails de l'explication moléculaire n'apportent rien, selon Kitcher, à la compréhension de la distribution des gènes en tant que processus-PS, mais encore ils ne l'expliquent pas vraiment, étant incapables de la représenter comme un processus-PS. En d'autres termes l'analyse moléculaire ne permet pas de regrouper sous un même type de phénomènes les mécanismes à l'œuvre au niveau de la cellule pendant la méiose. Quand bien même chaque division chromosomique peut en dernière analyse être décrite en termes moléculaires, aucune explication moléculaire ne permet d'unifier ces phénomènes, chose permise par le niveau d'explication cytologique.

La défense de l'anti-réductionnisme de la génétique classique à la génétique moléculaire repose donc, chez Kitcher, sur une inclusion de la cytologie dans sa définition de la génétique classique : la génétique classique n'est pas réduite à la génétique moléculaire parce que les explications *cytologiques* sont plus adéquates que les ex-

⁵¹Kitcher (1981, 1989) défend par ailleurs une thèse sur l'explication scientifique selon laquelle la valeur explicative d'une théorie est son pouvoir unificateur. La notion de schème de raisonnement est un des outils de cette thèse. Son argument anti-réductionniste est cependant intelligible sans qu'il soit nécessaire d'entrer en détail dans ses propos sur l'explication.

plications moléculaires. Incidemment, Kitcher semble donc identifier les schèmes de raisonnement propres à la génétique classique à des explications cytologiques. C'est sur ce point que je voudrais lui adresser une double critique.

2.2 Critique de la position de Philip Kitcher (1984)

J'ai insisté, dans la première section de ce chapitre, sur l'indépendance conceptuelle du mendélisme vis-à-vis de la théorie chromosomique de l'hérédité. En insistant sur la distinction entre génétique mendélienne et théorie chromosomique, je ne prétends aucunement nier l'importance des explications cytologiques dans le développement du mendélisme – le chapitre suivant en propose une analyse détaillée. Je ne prétends pas non plus, en refusant l'identification du mendélisme à une théorie cytologique, défendre une position réductionniste. Mon but, dans cette section, est de montrer, dans un premier temps (section 2.2.1), que Kitcher assimile à tort « schème de raisonnement » et « niveau d'explication », et de suggérer, dans un deuxième temps (2.2.2), que la caractérisation qu'il propose de la pratique de la génétique classique est insatisfaisante, en raison même du fait qu'elle est destinée à servir son argument anti-réductionniste.

2.2.1 Confusion entre schème de raisonnement et niveau d'explication

Comme je l'ai dit précédemment (section 2.1.1), le modèle classique de la réduction inter-théorique est formellement dépendant de l'idée d'une théorie présentée sous la forme d'un ensemble d'énoncés. Une théorie T_1 en réduit une autre T_2 si l'on peut déduire de T_1 la description de tous les phénomènes décrits (et expliqués) par T_2 . En caractérisant une théorie en termes de schèmes de raisonnement plutôt que sous la forme d'un ensemble hypothético-déductif d'énoncés, Kitcher cherche à montrer que la déductibilité logique d'un ensemble d'énoncés par un autre n'est pas un modèle adéquat de ce que l'on attend d'une explication scientifique. Expliquer – et comprendre – un type de phénomènes implique selon lui la mise en œuvre d'une certaine forme de raisonnement et non pas seulement la dérivation logique de la description du phénomène.

Cependant, quand il en appelle à sa caractérisation de la génétique pour attaquer la thèse réductionniste, Kitcher semble en revenir à une conception hypothético-déductive des théories : si l'explication cytologique est la plus adéquate, selon lui, c'est parce qu'il « y a un schème de raisonnement qui est appliqué pour *dériver la description*⁵² de l'assortiment des allèles et qui implique des prédicats qui caractérisent les cellules et leurs structures internes à une grande échelle. » Ainsi, *au sein*

⁵²Je souligne.

même du schème de raisonnement propre à la génétique classique (qui conduit à décrire la distribution des gènes dans les générations successives), interviennent, selon Kitcher, des *prémisses proprement cytologiques*. Je suggère donc que Kitcher assimile à tort deux thèses différentes. La première est celle selon laquelle la cytologie permet d'identifier un type de processus (la division des chromosomes pendant la méiose) qui appartient à la même classe (celle des processus-PS) que les processus de distribution des gènes – observés par expérience d'hybridation, et donc spécifiquement mendéliens. La seconde est celle selon laquelle l'explication cytologique *appartient* aux schèmes de raisonnement typiques de la génétique classique, auquel cas il n'y aurait plus de distinction à faire entre deux types de processus-PS (génétiques et cytologiques).

Cette assimilation grève l'argumentation de Kitcher à plusieurs titres. Une fois que la seconde thèse est acceptée, la première n'a plus vraiment lieu d'être : si les schèmes de raisonnement génétiques instancient des explications cytologiques, c'est-à-dire si les explications génétiques *sont* des explications cytologiques, l'affirmation selon laquelle la cytologie permet d'expliquer les phénomènes génétiques repose sur une pétition de principe. La question n'est plus alors celle de la réduction de la *génétique* classique à la génétique moléculaire, mais celle de la cytologie à la biologie moléculaire. Il semble pourtant que Kitcher ne prétende pas *identifier* les processus de distribution des gènes à des processus de division méiotique : affirmer, comme il le fait, que le niveau cytologique est le bon niveau d'explication des phénomènes génétiques revient à dire que les phénomènes à expliquer ne sont pas eux-mêmes initialement décrits au niveau cytologique.

Il convient donc, semble-t-il, de refuser la seconde thèse énoncée ci-dessus, c'est-à-dire l'identification des schèmes de raisonnement génétiques à la mise en œuvre d'une explication cytologique. Notons d'ailleurs que, dans les schèmes de raisonnement décrits par Kitcher, le niveau d'explication cytologique n'a pas une place bien définie : certes, dans l'ensemble des contraintes qui s'exercent sur la description de la transmission des gènes, on convoque des informations cytologiques, mais, au sein même du schème de raisonnement destiné à résoudre les problèmes de pedigree, ces informations n'ont pas de rôle *explicatif*. Plus précisément, ce n'est pas en tant qu'elles sont cytologiques qu'elles ont un rôle explicatif dans le cadre du schème de raisonnement.

Cela devient clair quand on couche ce raisonnement sous la forme d'un diagramme : les informations cytologiques jouent un rôle de contrainte disant quel type de dérivation est permise, mais rien, dans les règles de manipulation du diagramme, ne fait intervenir la considération des chromosomes. Prenons par exemple le diagramme de la figure 49 : la présence d'icônes indiquant l'emplacement des gènes responsables du daltonisme et de l'hémophilie sur le chromosome X permet de comprendre

la distribution de ces gènes à la troisième génération. Cependant, ces informations « cytologiques » pourraient très bien être données sans référence aux chromosomes : indiquer que les gènes responsables des caractères étudiés appartiennent à un même groupe de liaison et donner leur fréquence de recombinaison permettrait d'obtenir les mêmes résultats.⁵³

Le type de représentations adapté à une explication cytologique, comme je l'ai suggéré précédemment et comme j'y reviendrai au chapitre suivant, est celui des représentations *schématiques*, plutôt que diagrammatiques. En effet, la cytologie – tout comme, d'ailleurs, la biologie moléculaire – a pour objet le comportement d'entités concrètes (cellules, chromosomes, etc.) que l'on peut observer, par exemple, au microscope. Une explication cytologique consiste donc à faire référence à ces entités concrètes de manière essentielle, ce qui n'est pas le cas dans le raisonnement illustré par le diagramme de la figure 49, où le rôle des icônes représentant des chromosomes pourrait être joué par n'importe quel symbole contenant des informations qui elles-mêmes ne supposent pas l'hypothèse chromosomique. Même si l'emplacement des gènes étudiés explique certaines données numériques, la présence des chromosomes dans le diagramme est purement iconique ; ils ne fournissent pas d'explication mécanique *au sein du diagramme*.

Je suggère donc que la stratégie de Kitcher qui consiste à caractériser la génétique par les problèmes et les schèmes de raisonnement qu'elle met typiquement en œuvre pour les résoudre gagnerait, dans un premier temps du moins, à distinguer le niveau génétique du niveau cytologique, afin précisément d'évaluer la valeur explicative de ce dernier niveau. Ainsi, par exemple, il conviendrait de remplacer les diverses occurrences du terme « chromosome » par l'expression strictement mendélienne de « groupe de liaison ». Comme cela apparaîtra au chapitre suivant, non seulement cela rend compte plus fidèlement du développement historique de la génétique⁵⁴, mais encore cela permet de comprendre le rôle des hypothèses cytologiques dans le développement du mendélisme.

Ce n'est donc qu'en refusant la seconde thèse mentionnée ci-dessus – celle de l'in-

⁵³ Autrement dit, la cytologie ne sert à rien pour résoudre un problème de fréquence. Si l'on veut connaître, par exemple, la fréquence des yeux blancs dans une descendance de drosophiles, ou celle de l'hémophilie dans la famille de Victoria, il ne sert à rien de regarder les chromosomes de cette drosophile ou de Victoria avec un microscope. Merci à Patrice David pour cette remarque.

⁵⁴ En effet, si l'on prête attention aux débats des années 1920, certaines affirmations de Kitcher, comme la suivante, sont strictement fausses : « tout au long de l'histoire de la génétique classique, la théorie est orientée en direction de la réponse à des questions sur la distribution de caractéristiques dans les générations successives d'une généalogie, et elle propose de répondre à ces questions en utilisant *des probabilités de distributions de chromosomes* [je souligne] pour calculer les probabilités des génotypes de la descendance. » (Kitcher, 1984, p. 354)

clusion du niveau d'explication cytologique dans les schèmes de raisonnement mis en œuvre pour résoudre des problèmes de pedigree – que l'on peut aborder la question du réductionnisme. La thèse de Kitcher, une fois posée dans ces termes, devient la suivante : la génétique – définie comme étude des problèmes de pedigree – est réductible à la cytologie, mais non à la biologie moléculaire. Son argument, rappelons-le, consiste à affirmer que la cytologie permet d'identifier un type de processus cellulaires qui appartient, comme les phénomènes de distribution des gènes, à un type de processus-PS ; selon lui, la biologie moléculaire ne permet pas, en revanche, d'identifier un type homogène de processus-PS.

Ainsi formulée, cette thèse est très contestable, comme le montre Kenneth Waters (1990). En effet, en admettant que la cytologie offre un modèle mécanique satisfaisant pour expliquer la ségrégation et l'assortiment indépendant, Waters montre qu'elle ne fournit pas de meilleure explication que la biologie moléculaire pour des phénomènes génétiques comme ceux de liaison partielle.

Au niveau cytologique, on ne peut pas dire grand chose [du processus de crossing-over] sinon que les chromosomes homologues s'entremêlent parfois et échangent des segments pendant la division cellulaire. En revanche, peu de temps après que la structure en double hélice des chaînes génétiques fut comprise, des modèles moléculaires de crossing-over furent proposés. Le modèle basique de Holliday (Holliday, 1964), [voir figure 50] a été particulièrement fécond. [...] Notre compréhension de l'échange de segments entre des paires de gènes est considérablement élargie par notre connaissance de la structure moléculaire de ces chaînes. La biochimie de la recombinaison génétique est un domaine de recherche extrêmement actif et va conduire notre compréhension des processus mendéliens classiques au niveau moléculaire. » (Waters, 1990, p. 132)

Cet argument avancé par Waters dans une perspective réductionniste me permet, *a contrario*, de montrer que Kitcher n'est pas allé assez loin dans la perspective anti-réductionniste qui est la sienne. En affirmant la réductibilité de la génétique classique à la génétique moléculaire, Waters remet du même coup en question le fait que la cytologie fournisse une explication satisfaisante – que cette explication soit ou non réductible au niveau moléculaire – des phénomènes génétiques. En outre, la considération du modèle de Holliday (figure 50) met en évidence le fait que l'explication – qu'elle soit moléculaire ou cytologique (voir figure 39) – de phénomènes génétiques comme celui du crossing-over prend la forme d'une représentation schématique, comme je le suggérais précédemment. Les diagrammes de pedigree, quelle que soit la nature des icônes indiquant leurs règles de manipulation, ne constituent en aucun cas des explications mécaniques (cytologiques ou moléculaires, ou de toute autre sorte) des phénomènes génétiques.

Mon but, encore une fois, n'est pas de prendre part au débat sur le réduction-

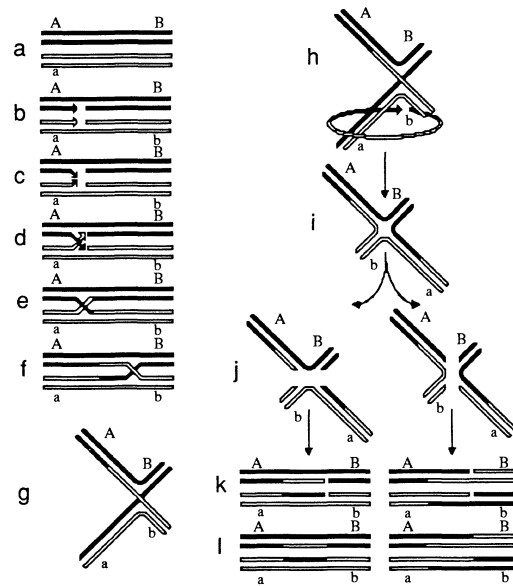


FIG. 50 – Le modèle de Holliday de la recombinaison génétique. (a) Deux doubles hélices homologues sont alignées. (b) Les deux brins + ou – sont coupés. (c) Les extrémités ainsi libérées quittent les brins complémentaires auxquels elles ont été reliées par un lien hydrogène. (d) Les extrémités libérées s'associent aux brins complémentaires de la double hélice complémentaire. (e) La ligation crée des doubles hélices partiellement hétéroduplexes. (f) Une migration du point d'embranchement se produit par la continuation du transfert du brin par deux chaînes polynuclotides impliquées dans le crossing-over. (g) La structure de Holliday présentée de manière étendue. (h) La rotation de la structure présentée en (g) donne la forme présentée en (i). La résolution de la structure présentée en (i) peut se produire de deux manières, selon le point de clivage enzymatique, ce qui donne les structures présentées dans (j), qui peuvent être représentées comme dans (k), et réappariées sous la forme présentée en (l). La figure est tirée de (Potter et Dressler, 1979, p. 970). L'explication est empruntée à (Suzuki *et al.*, 1986, p. 360). L'ensemble est tiré de (Waters, 1990).

nisme. Je suggère au contraire que la démarche de Kitcher et sa recherche d'une caractérisation non positiviste des théories et plus précisément de la génétique, est grevée par le fait qu'elle se situe dans le cadre du débat sur le réductionnisme. Alors que son but initial est de proposer une approche nouvelle des théories et des relations inter-théoriques (entre versions successives d'une même théorie et entre théories différentes), il concède finalement beaucoup à une conception hypothético-déductive des théories. Cela se reflète dans la manière même dont il propose de caractériser la génétique classique.

2.2.2 Les « problèmes de pedigree » ne sont pas le cœur du mendélisme

Je voudrais pour terminer comparer la manière dont Kitcher caractérise la génétique classique avec ma propre approche. Je prétends montrer deux choses : d'une part que sa caractérisation de la génétique est insatisfaisante, et d'autre part que ses défauts sont en partie dus aux limites de la définition même qu'il propose de la pratique d'une science.

J'ai montré, dans ce qui précède, que les schèmes de raisonnement propres à la génétique tels que Kitcher les décrit semblent assez fidèlement incarnés par les diagrammes de pedigree. Je voudrais à présent montrer que ces diagrammes et, conséquemment, les problèmes qu'ils présentent et servent à résoudre, ne constituent pas le cœur du mendélisme.

On peut, certes, utiliser un diagramme de cette sorte pour exprimer les lois de Mendel, comme dans la figure 44. Cependant, on peut également se servir d'une forme d'arbre généalogique du même type pour exprimer des hypothèses et mettre en œuvre un raisonnement non mendélien (comme c'est par exemple le cas dans l'arbre généalogique des Rougon-Macquart que j'ai placé en tête de ce chapitre). La forme de l'arbre généalogique précède largement la génétique mendélienne, dont je suggère que l'une des composantes théoriques fondamentales réside bien plutôt dans l'utilisation du symbolisme introduit par Mendel et dans l'utilisation des mathématiques combinatoires que dans le recours à ces diagrammes. Ils sont, certes, omniprésents dans les manuels et offrent une forme particulièrement commode pour présenter et résoudre des problèmes de pedigree. Mais les hypothèses authentiquement génétiques se manifestent dans l'utilisation des mathématiques combinatoires plus que dans celle d'une forme classiquement utilisée pour représenter l'hérédité, quelle qu'en soit la conception qu'on en ait.

À la rigueur, la présentation des problèmes génétiques sous la forme de diagrammes de pedigree est plutôt un héritage non mendélien de l'hérédité pensée comme « sommation de l'ensemble des influences de tous les ancêtres sur un individu ». Ces propos de Jean Gayon invitent à tirer une telle conclusion :

Dans le mendélisme, la question de l'origine des caractères est non pertinente : seule importe la structure génotypique présente, et les combinaisons qu'elle autorise. Aussi le pedigree n'a-t-il plus rien à dire de la nature de l'hérédité ; il n'intervient plus qu'à titre d'outil de mise en ordre des données. (Gayon, 1994)

Les diagrammes de pedigree offrent un cadre qui peut recevoir des éléments symboliques exprimant des hypothèses mendéliennes, mais qui pourraient servir – et ont servi – à exprimer des hypothèses très différentes sur l'hérédité. Ce qui est mendélien dans ces diagrammes, c'est l'usage de lettres ou autres symboles et icônes pour re-

présenter les facteurs ou les caractères discrets, sur lesquels on procède à des calculs combinatoires ; mais ils ne sont pas mendéliens en tant que diagrammes *de pedigree*.⁵⁵

Je suggère donc que les problèmes de pedigree, tels que Kitcher les définit, ne permettent pas vraiment de caractériser l'activité théorique propre au mendélisme. La génétique a, certes, pour l'objet l'étude de la transmission des facteurs héréditaires de génération en génération. Mais ce qui la caractérise au premier chef est, comme je l'ai suggéré au début de ce chapitre, la manière dont elle représente ces facteurs et son recours aux mathématiques combinatoires. L'inexactitude de la caractérisation de Kitcher devient patente dans quelques remarques qu'il ajoute après avoir présenté ce qu'il considère être le cœur de la pratique de la génétique classique :

La cœur de la théorie est la théorie de la *transmission des gènes*, la famille des schèmes de raisonnement dirigés vers les problèmes de pedigree. À cette théorie sont reliées d'autres sous-théories. La théorie de la *cartographie génétique* offre un schème de raisonnement qui aborde la question de l'emplacement relatif des loci sur les chromosomes. (Kitcher, 1984, pp. 357-358)

Non seulement Kitcher ne donne aucune définition de ce qu'il entend par « sous-théorie », mais encore il semble considérer que la cartographie génétique est l'application d'un schème de raisonnement différent de celui qui fait le cœur de la génétique classique. Je montrerai au contraire au chapitre suivant que cette technique représentationnelle peut être considérée comme une extension de première importance du type de représentation propre au mendélisme. En outre, on verra que l'analyse de la construction et de l'interprétation de ces cartes permet de comprendre et d'évaluer la nature explicative de l'hypothèse chromosomique et le rôle qu'elle a effectivement joué dans le développement de la génétique classique.

La recherche par Kitcher d'une « structure commune » aux « versions » successives de la génétique classique est une concession aux approches abstraites des théories scientifiques, qui l'empêche de rendre compte de la pratique effective de cette science. On peut ainsi faire à Kitcher une critique analogue à celle que Shapere (1964) fait à Kuhn (1962/1970), et que j'ai présentée au chapitre 6 (section 1.3.1) : en ayant recours à une notion somme toute assez abstraite et mal définie (la notion de schème de raisonnement, en tant qu'elle est une des composantes de la « pratique » d'une science), Kitcher ne passe pas la génétique par un tamis assez fin. Dans un souci pourtant anti-réductionniste, il manque ce qui fait la spécificité de l'activité

⁵⁵Voir (Gayon, 1994), qui compare les figurations « ancestraliste » et « mendélienne » des générations, et montre que, dans le contexte de l'hérédité ancestrale, les générations parentales convergent vers la génération filiale, présentant une « vision de l'hérédité comme ensemble d'influences dont le poids s'accumule dans l'hérédité », alors que la figuration mendélienne des pedigrees « descend vers les générations filiales (F_1 , F_2 , F_3) ».

théorique propre au mendélisme, quels que soient, ensuite, les supports matériels permettant d'expliquer les processus décrits au niveau génétique. Une fois rejetée la notion positiviste de théorie, il n'y a plus aucune raison de supposer l'existence d'une « structure » sous-jacente qui s'exprimerait dans des schèmes de raisonnement.

Plutôt que de reconstruire un schème de raisonnement abstrait différemment instantié au cours du temps, je suggère qu'une attention aux représentations concrètes et à la manière dont les différents généticiens les utilisent et les interprètent permet d'éclairer la nature de l'activité théorique propre au mendélisme et de comprendre le rôle des explications cytologiques mieux que ne le permet l'analyse de Kitcher. Dans cette perspective, les différentes « versions » de la génétique ne correspondront pas tant à différentes instantiations d'un schème abstrait de raisonnement qu'à différentes manières, *aussi bien diachroniques que synchroniques*, d'utiliser et de comprendre la méthode et le symbolisme mendéliens. En distinguant clairement entre le mendélisme et l'hypothèse chromosomique, il devient en effet possible de montrer que les différents types d'adhésions à cette dernière hypothèse déterminent différentes manières de comprendre et de pratiquer le mendélisme – différentes versions. C'est ce que l'étude de cas du prochain chapitre va chercher à montrer.

Conclusion

Soulignons, pour conclure, le contraste entre une approche des *théories* fondée sur la notion de réduction et ma propre caractérisation de l'*activité théorique*, en tant qu'elle repose sur la construction et l'utilisation d'un certain type de représentations. Pour reprendre une distinction introduite au début de ce travail, la compréhension d'une théorie se distingue de la capacité à opérer des déductions à partir de principes abstraits : elle a une double dimension computationnelle et représentationnelle. Comprendre la génétique mendélienne consiste ainsi à être en mesure de résoudre des problèmes au moyen du symbolisme qui lui est propre et à se représenter les phénomènes génétiques au moyen de ce symbolisme. De ce point de vue, comprendre ce qu'est un gène dans la perspective mendélienne et dans la perspective moléculaire sont clairement des opérations différentes.

Plutôt que de chercher à reconstruire la « structure » d'une théorie – que ce soit sous la forme d'un ensemble de lois, ou, comme Kitcher, sous la forme de « schèmes de raisonnement » – pour ensuite en analyser les relations avec d'autres théories, ou pour en dégager des versions successives, je suggère donc qu'une analyse des types de représentation utilisés dans l'enseignement, le développement et l'application des hypothèses théoriques permet de mettre en évidence d'autres moteurs importants du développement scientifique que la recherche de la réduction explicative (dont je ne

nie pas qu'elle en soit un moteur).

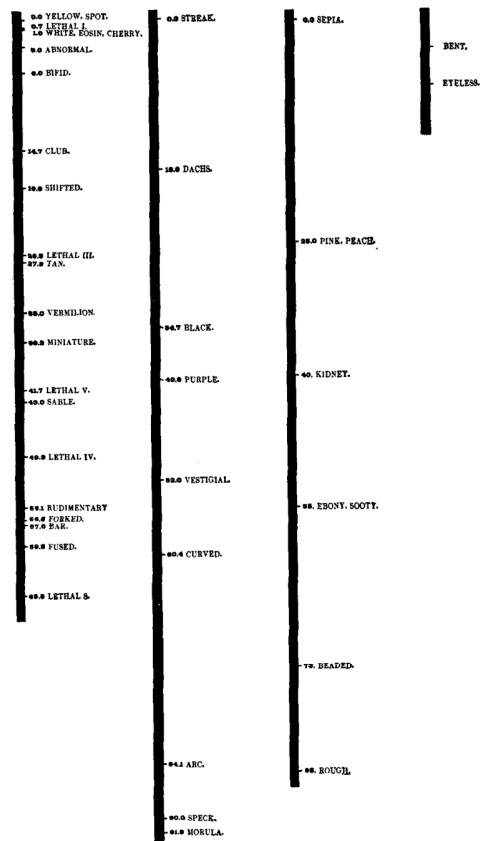
Ainsi, après avoir distingué entre le mendélisme « pur » et la théorie chromosomique en dégagant les types de représentations propres à chacun de ces domaines théoriques, je me propose, dans le prochain chapitre, d'étudier la manière dont ces deux domaines interagissent dans l'activité théorique des généticiens des années 1920. En effet, affirmer que le mendélisme est conceptuellement indépendant des différentes hypothèses possibles sur la nature matérielle des gènes ne revient pas à dire que les engagements théoriques des généticiens qui la développent ne jouent aucun rôle dans leur pratique et leur compréhension de la théorie.⁵⁶

Bien au contraire, le prochain chapitre est destiné à montrer la manière dont les options théoriques et les préférences méthodologiques des différents généticiens des années 1920 influent sur la façon dont ils construisent et interprètent les différentes représentations utilisées dans la pratique de la génétique. Cette influence devrait apparaître d'autant mieux qu'une distinction nette entre les principes de la génétique mendélienne et leur explication au niveau de la cellule a été tracée.

⁵⁶ Par exemple, la conviction de Hermann Muller que les gènes sont des particules matérielles dirige en grande partie son enquête, en collaboration avec le physicien Max Delbrück, après la découverte de l'effet mutagène des rayons X en 1927 (voir Morange, 2003, p. 26), sur la nature chimique des gènes, destinée à expliquer leur stabilité et leur capacité d'auto-réplication.

Chapitre 9

Étude de cas : la cartographie génétique (1913-1939)



(Frontispiece.)

Carte génétique des quatre chromosomes de *Drosophila Melanogaster*. Frontispice de Morgan *et al.*, 1915, *The Mechanism of Mendelian Heredity*.

La cartographie génétique est une technique mise en œuvre en 1913 par un élève de Morgan à l'Université de Columbia, Alfred Sturtevant. Elle consiste à représenter graphiquement l'emplacement relatif des gènes le long des chromosomes, sur la base de leur fréquence de recombinaison. Née de la suggestion de Morgan selon laquelle la fréquence de recombinaison de deux gènes appartenant à un même groupe de liaison est un indice de leur distance le long du chromosome, cette technique marque le début de la jonction progressive du mendélisme « pur » et de la théorie chromosomique de l'hérédité. À partir du milieu des années 1910, ce qui est « un des exemples les plus élégants de l'interaction du travail théorique et du travail expérimental dans l'histoire des sciences » (Wimsatt, 1987) devient le fil rouge des développements de la génétique par l'école de Morgan. La construction, l'interprétation et l'utilisation de ces cartes ont fait l'objet de plusieurs débats, qui offrent un laboratoire très utile pour une analyse de l'activité théorique dans la perspective qui est la mienne.

La cartographie génétique, et plus largement l'utilisation de représentations spatiales (picturales, diagrammatiques ou schématiques) en biologie ont fait l'objet de plusieurs études historiques et sociologiques, en particulier dans le courant des études sur la science [*science studies*].¹ Ces études adoptent la plupart du temps une perspective constructiviste sur l'activité scientifique, selon laquelle les faits scientifiques sont le résultat de processus de construction et de « négociation » sociologiquement déterminés. Mon analyse des représentations, si elle a beaucoup en commun avec cette étude des pratiques scientifiques centrée sur les « traces » produites par les théoriciens et les dispositifs expérimentaux, vise cependant à insister sur l'importance *théorique* de ces représentations : mon but, contrairement à Kohler (1994), n'est pas d'analyser le processus de « construction » de la drosophile, mais bien de comprendre la manière dont les engagements théoriques et les contraintes liées à la prédiction et à l'explication des phénomènes génétiques interagissent dans la construction et l'interprétation de certains types de représentations, que je considère comme des représentations *théoriques*.

De ce point de vue, ma perspective est plus proche de celle qu'adoptent Griesemer et Wimsatt (1989) dans leur étude de l'évolution des diagrammes de Weismann, qu'ils considèrent comme des entités *conceptuelles*. Plutôt que de procéder à une étude des changements de signification du concept de gène, par exemple, j'isole comme eux des types de représentations dont certains traits concrets – par opposition à une histoire conceptuelle dématérialisée – sont caractéristiques de certains domaines théoriques – mendélisme d'une part et cytologie d'autre part. Une des différences majeures entre leur approche et la mienne est cependant que mon analyse des *versions* du mendé-

¹Voir en particulier (Rheinberger et Gaudillière, 2004a,b; Kohler, 1994).

lisme, révélées dans la diversité des approches possibles de la cartographie, est plus *synchronique* que *diachronique* : je n'étudie pas les différentes formes sous lesquelles on retrouve le mendélisme à des époques et dans des contextes différents, mais j'analyse comment, à une période donnée, différentes attitudes épistémiques vis-à-vis des composantes théoriques de la génétique classique présentées au chapitre précédent se manifestent dans la manière dont les théoriciens construisent et interprètent les cartes génétiques.

Le fait que ces différentes attitudes aient parfois fait l'objet de débats explicites présente un grand avantage méthodologique. Il évite en effet de se livrer à une reconstruction historiquement invérifiable des conceptions des différents acteurs, sur la base de ce qui est « implicitement » contenu dans les représentations qu'ils utilisent et plus généralement dans leur propos.²

Après avoir, section 1, rappelé le double contexte théorique dans lequel la cartographie génétique a été inventée, je présente, à la section 2, la technique cartographique mise au point par Sturtevant (1913a). Les sections 3 et 4 exposent brièvement les développements ultérieurs de la cartographie dans les années 1920 et les étapes qui conduisent à la confirmation de l'hypothèse chromosomique en 1934. La section 5 présente les critiques qui ont été faites à cette technique et à l'hypothèse du crossing-over par trois généticiens contemporains de l'école de Morgan. L'analyse de ces débats à la section 6, conduite au moyen des outils d'analyse des représentations élaborés au chapitre 7, cherche à mettre en évidence la manière dont les engagements de chacun en faveur d'une hypothèse explicative au niveau cytologique guide sa construction et son interprétation des cartes, que je propose de décrire comme l'expression de sa *version* du mendélisme.

1 Préhistoire de la cartographie

Avant de présenter la technique de cartographie génétique proprement dite, il convient de dire un mot des découvertes et des problèmes qui précèdent immédiatement son élaboration en 1913 par Sturtevant. J'ai insisté, dans le chapitre précédent, sur la distinction entre le mendélisme « pur » et la théorie chromosomique de l'hérédité. J'ai montré, en particulier, que les activités théoriques propres à chacun de ces

²Cette notion d'« implicite » est très présente dans l'histoire de la génétique mendélienne de Lindley Darden (1991), qui insiste sur l'importance du symbolisme et des engagements théoriques véhiculés par l'utilisation de certains types de représentations. Tout en empruntant des éléments à son livre et en proposant aussi une approche du mendélisme par l'étude des symbolismes qui lui sont propres, je cherche autant que possible à me concentrer sur les pratiques objectives et les propos explicites des généticiens.

domaines se caractérisent par l'usage de représentations de types différents, symptomatiques de niveaux d'explication différents. Rappelons-les brièvement.

Le *mendélisme* repose sur l'utilisation des mathématiques combinatoires appliquées à des symboles représentant les gènes sans que soit faite aucune hypothèse sur leur nature. Ce symbolisme peut être étendu à des diagrammes comme les carrés de Punnett facilitant le raisonnement probabiliste. La méthode mendélienne consiste en une analyse quantitative de données obtenues par expériences de croisement (voir chapitre 8, section 1.1).

La *cytologie* est l'étude de la constitution et des processus cellulaires. Elle repose sur la manipulation de représentations picturales obtenues au microscope optique et sur la construction de schémas à partir de ces images (voir chapitre 8, section 1.4).

La cartographie génétique doit sa naissance à la rencontre de ces deux domaines, et plus particulièrement à l'hypothèse du crossing-over avancée par Morgan en 1912. Dans cette section, je reviens sur le double contexte théorique dans lequel la cartographie génétique a été inventée.

1.1 Le phénomène de liaison partielle

À partir de 1905, plusieurs exceptions à la loi d'assortiment indépendant sont observées au cours d'expériences de croisement génétique : certains facteurs tendent à être hérités ensemble, sans que cette liaison soit systématique (auquel cas on pourrait considérer les caractères en question comme étant causés par un seul allèle) ; on parle alors de « liaison partielle ». Les premiers à constater de telles anomalies sont Bateson, Punnett et Saunders, lors d'expériences de croisement sur le pois *Lathyrus odoratus* (Bateson *et al.*, 1905).

La couleur des fleurs d'une part et la forme des grains de pollen de l'autre ségrègent normalement suivant la première loi de Mendel : à la deuxième génération d'hybrides, on trouve une proportion de trois fleurs de couleur violette (A) pour une de couleur rouge (a) et une proportion de trois fleurs aux grains de pollen allongés (B) pour une aux grains de pollen ronds (b). Cependant, lors des croisements dihybrides impliquant les deux caractères, on ne trouve pas la proportion attendue de $9AB : 3Ab : 3aB : 1ab$. Les résultats rapportés par Bateson *et al.* (1905) sont les suivants : $1528AB : 106Ab : 117aB : 381ab$.

Les caractères A et B d'une part, et a et b d'autre part tendent à être hérités ensemble : leur fréquence d'association révèle que l'assortiment des facteurs responsables de la couleur de la fleur et de la forme des grains de pollen ne ségrègent pas indépendamment.

Au cours de la ségrégation, les aspects de la structure ou de la constitution physiologique sont traitées comme des *unités* par les divisions cellulaires au cours desquelles les germes sont formés. Ces caractères qui ségrègent sont, selon l'état actuel de nos connaissances, toujours constitués par paires. [...] Parfois, cependant, on constate la liaison ou le *couplage* entre des caractères distincts. Quand de tels couplages sont complets, les deux caractères peuvent bien sûr être traités comme un seul allélomorphe. [...] Mais, au delà de ce phénomène simple de couplage complet, on sait aujourd'hui que les ratios habituels sont susceptibles d'être modifiés par un *couplage partiel* entre des caractères distincts. (Bateson, 1906)

Insistons sur un point crucial : cette anomalie est découverte dans le cadre d'expériences de croisement mendéliennes. Elle n'a aucun lien avec l'hypothèse chromosomique, dont Bateson et Punnett resteront très longtemps de fervents opposants. L'explication qu'ils proposent dans un premier temps est la suivante (Bateson *et al.*, 1905, p. 141) : les gamètes contenant les différentes combinaisons de facteurs ne sont pas en nombre égal, la production gamétique de la première génération d'hybrides étant d'après leur hypothèse approximativement de $7AB + 1Ab + 1aB + 7ab$ au lieu d'être en nombre égal ($1AB + 1Ab + 1aB + 1ab$). L'anomalie, selon eux, provient du couplage des facteurs dominants. La suite de leurs recherches les conduit à trouver des cas un peu différents, qu'ils expliquent par la réciproque du couplage (ou attraction) : la répulsion.

[...] la conclusion à laquelle nous sommes arrivés est que si A , a , et B , b , sont deux paires d'allélomorphes sujettes au couplage et à la répulsion, les facteurs A et B vont se repousser mutuellement au cours de la formation des gamètes du double hétérozygote résultant de l'union $Ab \times aB$, mais vont être couplés au cours de la formation des gamètes du double hétérozygote résultant de l'union $AB \times ab$. L'hétérozygote F_1 est visiblement identique dans les deux cas, mais ses descendants présentent la distinction. Nous n'avons pour l'heure pas de conjecture probable à proposer concernant la nature essentielle de cette distinction, et tout ce qui peut être dit est que, dans ces cas spéciaux, la distribution des caractères dans l'hétérozygote est affectée par la distribution dans les parents purs. (Bateson et Punnett, 1911b, p. 216)

Cette explication est à comprendre dans le cadre de la théorie de la « présence-absence » de Bateson³, destinée à expliquer la dominance : la dominance est due à la présence du facteur, et la récessivité à son absence. A et B désignent la présence du caractère dominant, et a et b son absence. Par conséquent, si les facteurs A et B sont plus fréquemment présents, on dira qu'ils sont couplés (ou s'attirent). Dans le

³Voir (Carlson, 1967, chap. 8).

cas contraire, ils se repoussent.⁴

En 1906, Doncaster et Raynor constatent, chez le lépidoptère *Abraxas*, une anomalie d'une autre sorte : quand une femelle *laticolor* (de couleur pâle) est croisée avec un mâle *grossulariate* (de couleur sombre), tous les hybrides de la première génération sont *grossulariate*. À la deuxième génération d'hybride, une anomalie se présente : on ne trouve aucun mâle *laticolor*. Doncaster cherche à expliquer cette anomalie en termes de couplage entre le sexe féminin (interprété comme un facteur mendélien) et le facteur responsable du *laticolor*. Bateson et Punnett (1908) amendent cette explication et proposent celle, nécessitant des hypothèses moins nombreuses, d'une répulsion entre le facteur du sexe féminin et le facteur *grossulariate*.⁵

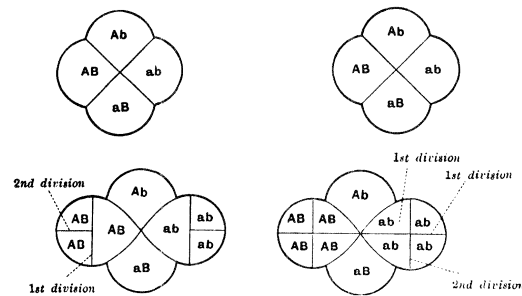


FIG. 51 — Schéma présentant l'hypothèse de reduplication. Bateson et Punnett expliquent les ratios non mendéliens obtenus au cours des croisements en supposant une multiplication différentielle des cellules (à gauche, $3AB : 3ab : 1Ab : 1ab$; à droite, $4AB : 4ab : 1Ab : 1aB$). Ces reduplications devaient être mitotiques, et non méiotiques; elles étaient, en outre, restreintes aux cellules germinales (Punnett, 1913).

Quelques années plus tard, (Bateson et Punnett, 1911a) proposent une explication mécanique aux phénomènes de couplage et de répulsion⁶, celle de la reduplica-

⁴Lindley Darden (1991, p. 122) remarque que la terminologie du couplage (ou attraction) et de la répulsion reflète la prédilection de Bateson pour les analogies physiques : il fait souvent référence à des ondes, des vortex et des charges (voir par exemple Bateson, 1913, chap. 2 et 3). Pour l'évolution intellectuelle de Bateson, voir (Coleman, 1970; Darden, 1977; Cock, 1983).

⁵Pour plus de détails sur ces débats, voir (Morgan, 1910a, p. 486-488).

⁶Les termes de couplage et de répulsion étaient en effet uniquement destinés à décrire l'association des facteurs présents dans les cellules germinales : « Au vu de ce que nous savons, les termes “couplage” et “répulsion” sont mal choisis. “Couplage” a d'abord été introduit pour dénoter l'association de facteurs spéciaux, et “répulsion” pour décrire la dissociation de facteurs spéciaux. À présent que l'on sait que ces phénomènes sont causés non par l'association ou la dissociation, mais par le développement de certaines cellules en excès, ces expressions doivent être abandonnées. Les termes indicatifs de multiplication différentielle ou de prolifération sont plus adaptés. Au point où nous en sommes, nous hésitons à suggérer ces termes, mais les divers systèmes peuvent être désignés comme des exemples de *reduplication*, quels que soient les moyens par lesquels la composition numérique

tion différenciée des cellules.⁷ Cette hypothèse permet de comprendre la production de différents types de gamètes en nombre inégal. Bateson et Punnett proposent ainsi d'expliquer les exceptions aux proportions normales comme le résultat d'une duplication différenciée de certaines cellules germinales (voir figure 51).

1.2 Morgan et les caractères liés au sexe

En 1910, au moment où il commence, à l'université de Columbia, ses travaux sur la mouche du vinaigre *Drosophila melanogaster*⁸ en collaboration avec deux de ses étudiants, Alfred Sturtevant et Calvin Bridges, bientôt rejoints par Hermann Muller, qui fait ses études au laboratoire du cytologiste Wilson, Morgan est toujours un opposant de la théorie chromosomique (Morgan, 1910a). Les premiers travaux sur la drosophile ont pour objet la recherche de mutations ou variations discontinues telles que de Vries en avait trouvées chez *Oenothera lamarckiana* (onagre) dès 1900, dans le cadre de l'élaboration de sa théorie de la mutation.⁹ C'est la découverte d'un mâle drosophile aux yeux blancs (figure 52) au sein d'une population de mouches aux yeux rouges (couleur normale), qui marque le début d'une série d'expériences qui conduisent Morgan à adopter l'hypothèse chromosomique.¹⁰



FIG. 52 – Drosophiles : mutant aux yeux blancs, et individu normal (yeux rouges).

des gamètes est produite. » (Bateson et Punnett, 1911a)

⁷Voir (Carlson, 1967, p. 49-52).

⁸Pour une histoire du groupe de la drosophile et en particulier du choix et de la « construction » de cet organisme modèle, voir (Kohler, 1994).

⁹Pour une histoire détaillée, voir (Dunn, 1965) et (Carlson, 1967, chap. 6).

¹⁰Pour une étude détaillée de l'évolution intellectuelle de Morgan et de sa conversion à la théorie chromosomique, voir (Carlson, 1967) et (Allen, 1978).

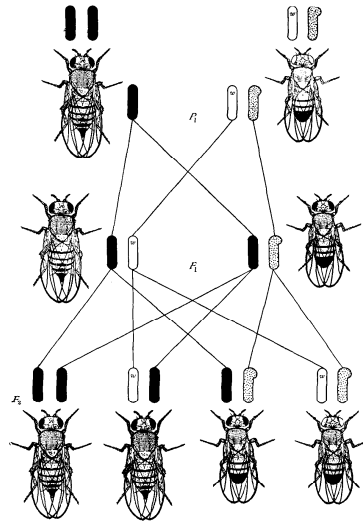


FIG. 53 – Représentation diagrammatique des proportions obtenues lors des expériences de croisement d'un mâle aux yeux blancs et d'une femelle aux yeux rouges. Les deux chromosomes X de la femelle de la génération parentale P_1 portent le gène pour la couleur rouge. Le gène responsable de la couleur blanche w est porté par le chromosome X du mâle, représenté en blanc. Le chromosome Y est moucheté. À la première génération d'hybrides F_1 , tous les individus, mâles et femelles, ont les yeux rouges, ce qui s'explique par la dominance du facteur pour les yeux rouges. À la génération F_2 , résultant du croisement des individus de la génération F_1 , tous hétérozygotes, on obtient bien trois individus aux yeux rouges pour un aux yeux blancs. En revanche, on ne trouve pas de femelle aux yeux blancs et, parmi les individus aux yeux rouges, la proportion est de deux femelles pour un mâle. (Ce diagramme est tiré de (Morgan, 1928, p. 60) ; à l'époque des expériences de Morgan (1910b), le sexe est considéré comme un facteur mendélien, et on ne parle pas encore de chromosome Y ; cette représentation est donc légèrement anachronique.)

Comme le résume la figure 53, le croisement du mâle aux yeux blancs avec des femelles normales (aux yeux rouges) ne donne, à la première génération d'hybrides, que des individus aux yeux rouges. La deuxième génération d'hybrides présente la proportion attendue de trois individus aux yeux rouges pour un individu aux yeux blancs avec, cependant, un phénomène anormal : tous les individus aux yeux blancs sont des mâles et, parmi les individus aux yeux rouges, la proportion est de deux femelles pour un mâle. Un deuxième croisement des mâles aux yeux blancs avec des femelles aux yeux rouges de la première génération d'hybride (F_1) donne une proportion égale d'individus aux yeux rouges et aux yeux blancs, des deux sexes. Morgan propose, en 1910, une explication qui ne fait pas intervenir le chromosome sexuel, en considérant le sexe comme un facteur mendélien (Morgan, 1910b). Pour un exposé détaillé de ces expériences sur le mutant « yeux blancs », voir l'article de Morgan (1910b) et (Carlson, 1967, chap. 6, p. 43-45). Morgan, par ailleurs, est particulièrement intrigué

par le fait que le cas du mutant yeux blancs chez la drosophile présente le phénomène inverse de l'anomalie détectée chez *Abraxas*, où le facteur récessif *laticolor* était lié au sexe féminin :

La considération la plus importante à tirer de ces résultats est qu'en tous points ils fournissent des résultats inverses de ceux obtenus par *Abraxas* dans les expériences de Punnett et Raynor. [...] Un autre fait significatif est le suivant : l'analyse des résultats montre que la femelle sauvage *Abraxas grossulariate* est hétérozygote pour ces deux caractères. (Morgan, 1910b)

Au cours de l'année 1910, Morgan isole d'autres mutations chez la drosophile manifestement liées au sexe, comme la couleur jaune du corps et la taille miniature des ailes. Jusqu'ici, la liaison de ces facteurs au sexe n'implique pas, pour Morgan, qu'ils soient situés sur le chromosome *X* identifié comme le chromosome sexuel par les cytologistes. Cependant, il apparaît que *les différents caractères liés au sexe sont eux-mêmes partiellement liés les uns aux autres* : on observe une recombinaison de facteurs lors de croisements d'individus aux yeux blancs avec des individus aux ailes miniatures et d'individus aux yeux blancs avec des individus au corps jaune. La rencontre des phénomènes de liaison partielle d'une part et de liaison au sexe d'autre part a joué un rôle déterminant dans l'adoption par Morgan de la théorie chromosomique de l'hérédité. En effet, l'explication qu'il donne en 1911 (Morgan, 1911c) de ces phénomènes situe les facteurs mendéliens sur les chromosomes.

1.3 L'hypothèse du crossing-over

Les recherches contemporaines des cytologistes sur les chromosomes sexuels des années 1905-1906¹¹ suggèrent une hypothèse satisfaisante. Comme je l'ai mentionné au chapitre précédent, la découverte qui va jouer un rôle crucial dans l'élaboration par Morgan de la théorie dite du « crossing-over » et son adoption définitive de l'hypothèse chromosomique date de 1909 : le cytologiste Janssens (1909) observe, au cours d'une des premières phases de la méiose, des entrecroisements entre chromosomes homologues, qu'il appelle « chiasmata » (voir figure 54). Il énonce la théorie de la « chiasmatypie », selon laquelle, au cours de ces entrecroisements, les chromosomes homologues échangent des portions (voir chapitre précédent, section 1.3.2).

De cette observation, Morgan tire un ensemble d'hypothèses, propres à apporter une explication mécanique simple des phénomènes de couplage partiel découverts par Bateson. Il suggère de considérer ces phénomènes comme le résultat d'un échange ordonné de facteurs entre groupes de liaison homologues, que Morgan et Cattell (1912) appellent pour la première fois « crossing-over ». Cet échange de facteurs pourrait

¹¹Voir (Dunn, 1965).

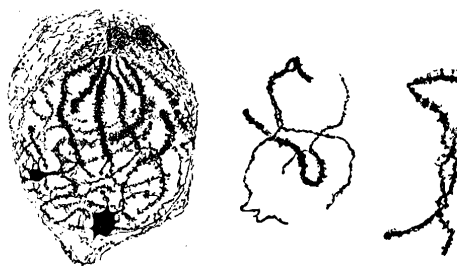


FIG. 54 – Entrecroisements de chromosomes pendant la méiose (Janssens, 1909).

s'expliquer, au niveau cytologique, par un échange de portions entre chromosomes homologues pendant la méiose.

Au lieu d'attractions, de répulsions et d'ordre de précédence, et des systèmes compliqués de couplage, je me hasarde à suggérer une explication plus simple fondée sur les résultats concernant l'hérédité des facteurs responsables de la couleur des yeux, de la couleur du corps, des mutations des ailes et du sexe féminin chez *Drosophila*. Si les matériaux qui représentent ces facteurs sont contenus dans les chromosomes, et si ces facteurs qui sont « couplés » se trouvent *à proximité sur une même ligne*¹², alors, quand des paires parentes (chez les hétérozygotes) sont conjuguées, les régions apparentées vont se trouver face à face. Il y a de bonnes raisons de suggérer que [...] les chromosomes homologues s'entrelacent, et que, quand les chromosomes se séparent, la séparation est faite sur un même plan, comme le soutient Janssens. En conséquence, les matériaux d'origine sont plus susceptibles, pour les courtes distances, de se retrouver du même côté de la séparation, alors que les régions plus éloignées sont aussi susceptibles de se retrouver du même côté que du côté opposé. En conséquence, on observe le couplage de certains caractères, et très peu de cas de couplage d'autres caractères ; la différence dépend de la distance linéaire qui sépare les matériaux chromosomiques qui représentent les facteurs. (Morgan, 1911c)

Notons que Morgan reste très vague sur la nature des facteurs héréditaires et sur la manière dont ils sont attachés aux chromosomes : pour reprendre ses termes, ils sont « représentés » par des morceaux de chromosomes. Le modèle du collier de perles (figure 55), proposé quelques années plus tard par Morgan *et al.* (1915), n'en dit pas plus sur la nature du lien entre chromosomes et gènes. Dans ce modèle, les gènes sont représentés comme des perles alignées le long d'un fil, sans qu'il soit question de représenter par là la structure réelle du chromosome (voir chapitre 8, section 1.3.2). La seule hypothèse que ce modèle prétend vraiment exprimer, et qui sous-tend celle du crossing-over, est celle de l'*arrangement linéaire* des gènes.

¹²Je souligne.

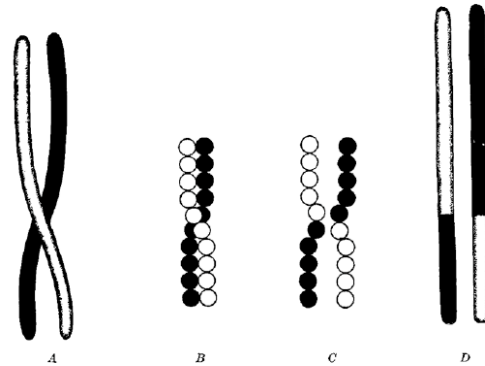


FIG. 55 – Le modèle du « collier de perles », présenté par Morgan *et al.* (1915, p. 60), est une explication mécanique du crossing-over. Les gènes sont représentés comme des perles alignées le long des chromosomes. Les chromosomes sont représentés comme composés d'une unique chromatide. À gauche (A), les chromosomes parentaux ; au centre (B et C), crossing-over (échange de « perles ») dû à un chiasma cytologique ; à droite (D), chromosomes recombinants. Lorsqu'un échange de brin se produit, plus la distance qui sépare deux facteurs est grande, plus ceux-ci ont de chance d'être recombinés au cours de la formation des gamètes.

Pour bien comprendre l'importance de cette idée de linéarité, et la valeur explicative de l'hypothèse du crossing-over avancée par Morgan, il convient de préciser un certain nombre de caractéristiques du phénomène de liaison partielle. Pour le croisement dihybride concernant deux caractères portés par les deux facteurs A et B , un cas de liaison totale donnerait lors du croisement de deux parents homozygotes pour les deux caractères, l'un dominant l'autre récessif ($AABB$ et $aabb$), dans les cellules germinales produites, 50% de AB , 50% de ab , mais aucun Ab ni aB . La loi d'assortiment indépendant prévoit, elle, 25% de AB , 25% de ab , 25% de aB et 25% de Ab . La liaison partielle se caractérise par le fait que les facteurs A et B d'une part et a et b de l'autre se trouvent liés dans $50 - r\%$ des cas, tandis que les facteurs A et b et a et B se trouvent liés dans $r\%$ de cas.

Or on trouve que cette proportion r est

- constante pour toute paire de facteurs donnés ;
- différente pour différentes paires de facteurs ;
- indépendante de la combinaison initiale des parents (c'est-à-dire que pour des parents $AAbb$ et $aaBB$, on trouve le résultat suivant au niveau des types de gamètes : $r\%$ de AB , $50 - r\%$ de Ab , $50 - r\%$ de aB , et $r\%$ de ab .)

Si l'on appelle R la proportion $2r$, on dit qu'il y a recombinaison dans $R\%$ des cas.

L'hypothèse du crossing-over, jointe à son explication cytologique, se décompose dans les hypothèses suivantes :

- les gènes sont ordonnés linéairement sur les chromosomes (hypothèse de linéa-

AB	Ab	aB	ab
50	0	0	50
25	25	25	25
50-r	r	r	50-r

TAB. 1 – Le phénomène de liaison partielle

rité) ;

- lors de la méiose, les chromosomes homologues qui s'entrecroisent échangent des parties (hypothèse de la chiasmatypie) ;
- donc la fréquence d'association ou de séparation des facteurs – dont on a vu qu'elle était constante pour toute paire de facteurs donnée et différente pour différentes paires de facteurs – est un produit de la fréquence de cassure entre différents loci¹³ du chromosome, et est donc *fonction de la distance qui sépare les facteurs le long du chromosome* (en vertu de l'hypothèse de linéarité).

Une remarque de vocabulaire s'impose : ici, la théorie du crossing-over est présentée dans les termes de l'hypothèse selon laquelle les gènes sont localisés sur les chromosomes. À strictement parler, cependant, la théorie du crossing-over dit simplement que les gènes d'un même groupe de liaison sont ordonnés linéairement, et que des échanges de portions se produisent entre groupes homologues. L'observation cytologique de Janssens, ainsi que la thèse de détermination chromosomique du sexe, étant à l'origine de la formulation de la théorie du crossing-over par Morgan, et le terme de « crossing-over » ayant lui-même une connotation mécaniste, il arrive que l'on parle de « théorie du crossing-over » pour désigner l'hypothèse chromosomique elle-même. Cependant, comme on va le voir, non seulement ces deux hypothèses sont conceptuellement indépendantes, mais encore plusieurs généticiens ont effectivement adopté chacune d'entre elles en refusant l'autre.

L'hypothèse chromosomique fournit ainsi une explication mécanique (voir figure 55) du phénomène de liaison partielle, observé au niveau génétique.

Les résultats sont simplement le résultat mécanique de l'emplacement des particules matérielles sur les chromosomes, et de la méthode d'union des chromosomes homologues, et les proportions qui en ressortent ne sont pas tant l'expression d'un système numérique que de l'emplacement relatif des facteurs sur le chromosome. (Morgan, 1911c, p. 384)

Morgan (1911b) explique ainsi sa conversion :

Pendant plusieurs années, il m'a semblé que l'hypothèse chromosomique, en tant

¹³Un *locus* correspond à l'emplacement précis d'un gène sur un chromosome.

que telle, ne pouvait pas être utilisée pour expliquer les résultats mendéliens sous la forme présentée par Sutton, parce que, si elle était vraie, il n'aurait pas pu y avoir plus de paires mendéliennes dans une espèce donnée que le nombre chromosomes présents dans cette espèce. Même si cette objection pouvait être évitée [...], une objection plus grave demeurait : avec un petit nombre de chromosomes, de nombreux caractères auraient dû mendéliser [*Mendelize*] ensemble, et très peu de cas de cette sorte sont connus. De Vries¹⁴ fut le premier, je crois, à mettre en évidence que cette objection pouvait être évitée si les gènes étaient contenus dans des corps plus petits qui pouvaient passer entre les paires de chromosomes. Dans le cas de l'hérédité de deux caractères limités au sexe dans le même animal, on a une vérification expérimentale de cette hypothèse. (Morgan, 1911b, p. 77-78)

Comme on va le voir, la technique de cartographie génétique, mise au point par Alfred Sturtevant, est née de la suggestion de Morgan selon laquelle la fréquence d'association (ou de séparation) des facteurs est une fonction de leur distance le long du chromosome.

2 Sturtevant et la première carte génétique (1913)

Alfred Sturtevant est encore étudiant sous la direction de Morgan quand il construit, en 1913, la première carte génétique, présentant du même coup la technique de cartographie conçue à partir de l'hypothèse de Morgan. Dès son article de 1913, la plupart des problèmes explorés par le groupe de Morgan dans les années qui suivent se font jour.

2.1 La technique d'Alfred Sturtevant

Sturtevant (1913a) transforme l'hypothèse de Morgan en une technique quantitative qui consiste à utiliser la fréquence des crossing-over observés (par la recombinaison des facteurs) lors d'expériences de croisement comme une donnée numérique permettant de produire l'ordre relatif des gènes sur le chromosome. La technique de cartographie est donc un schéma quantitatif qui convertit les fréquences de crossing-over en un graphe représentant les emplacements relatifs des gènes disposés linéairement sur le chromosome.

Si cette hypothèse [celle de Morgan] est correcte, il semble que la proportion de « crossing-over » puisse être utilisée comme un indice de la distance entre deux

¹⁴En 1903, Hugo de Vries avait fait l'hypothèse d'un échange de facteurs portés par les chromosomes, mais elle n'était soutenue par aucune observation empirique et n'avait pas de valeur prédictive (voir Carlson, 1967, p. 48).

facteurs. Ainsi, en déterminant la distance (ainsi définie) entre A et B et entre B et C , on doit pouvoir prédire AC . En effet, si les proportions de crossing-over *représentent réellement la distance*¹⁵, AC doit correspondre approximativement soit à la somme soit à la différence entre AB et BC , et à aucune valeur intermédiaire. (Sturtevant, 1913a, p. 3)

Ainsi, les degrés de séparation entre A et B et entre B et C doivent permettre de prédire le degré de séparation entre A et C , selon l'égalité $R(AB) + R(BC) = R(AC)$.¹⁶ En prenant plusieurs paires de facteurs liés au même chromosome, on doit pouvoir progressivement construire la carte de leur emplacement relatif.

Sturtevant (1913a) construit donc la carte du chromosome X de la drosophile présentant l'emplacement relatif des six facteurs liés au sexe identifiés par Morgan depuis 1911. Ces facteurs (désignés par les lettres B , C , O , P , R et M) concernent la couleur du corps, des yeux et la taille des ailes de la drosophile (Sturtevant, 1913a, p. 3-4). Par l'observation des recombinaisons des facteurs liés au sexe chez les mâles drosophile – en effet, chez les femelles, en raison de la présence de deux chromosomes X , la présence d'un facteur récessif lié au sexe sur l'un des chromosomes est inobservable au niveau phénotypique (Sturtevant, 1913a, p. 4) –, il calcule les « distances » relatives entre chacun de ces facteurs.

L'unité de « distance » est donnée par une portion de chromosome d'une longueur telle que, en moyenne, un crossing-over s'y produit pour 100 gamètes formés. Autrement dit, le pourcentage de crossing-over est un indice de la distance.

La figure 56 présente les proportions et pourcentages de crossing-over (*i.e.* les distances) entre chaque paire de facteurs. De ces résultats numériques, Sturtevant peut tirer une représentation graphique (figure 57) des distances relatives des facteurs au facteur B (responsable de la couleur du corps).

Sturtevant précise que la construction de la carte s'effectue en s'appuyant en priorité sur les *petites distances*, c'est-à-dire les paires de facteurs fortement liés (qui recombinent rarement). Il constate en effet que les prédictions sont beaucoup moins précises pour les grandes distances. Alors que l'hypothèse de linéarité semble confirmée par l'additivité des (relativement)¹⁷ petites distances, les proportions de crossing-over entre deux facteurs éloignés comme B et M prédites par l'addition des distances entre BP et PM par exemple ne correspondent pas au nombre de recombinaisons effectivement observées (voir figure 58).

¹⁵ Je souligne.

¹⁶ R désigne la fréquence de recombinaison (voir page 554).

¹⁷ Voir plus bas sur l'effet d'interférence.

Factors concerned	Proportion of crossovers	Percent of crossovers
BCO	$\frac{193}{16287}$	1.2
BO	$\frac{2}{373}$	0.5
BP	$\frac{1464}{4551}$	32.2
BR	$\frac{115}{324}$	35.5
BM	$\frac{260}{693}$	37.6
COP	$\frac{224}{748}$	30.0
COR	$\frac{1643}{4749}$	34.6
COM	$\frac{76}{161}$	47.2
OP	$\frac{247}{836}$	29.4
OR	$\frac{183}{538}$	34.0
OM	$\frac{218}{404}$	54.0
CR	$\frac{236}{829}$	28.5
CM	$\frac{112}{333}$	33.6
B(C, O)	$\frac{214}{21736}$	1.0
(C, O) P	$\frac{471}{1584}$	29.7
(C, O) R	$\frac{2062}{6116}$	33.7
(C, O) M	$\frac{406}{898}$	45.2
PR	$\frac{17}{573}$	3.0
PM	$\frac{109}{405}$	26.9

FIG. 56 – Tableau des résultats numériques. (Sturtevant, 1913a)

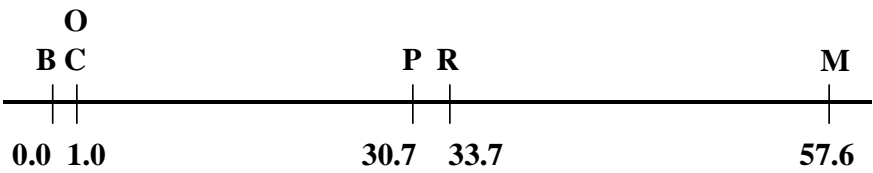


FIG. 57 – Carte génétique du chromosome X. (Sturtevant, 1913a)

Factors	Calculated distance	Observed per cent of crossovers
BP	30.7	32.2
BR	33.7	35.5
BM	57.6	37.6
(C, O) R	32.7	33.7
(C, O) M	56.6	45.2

FIG. 58 – Exceptions à l'additivité pour les grandes distances (Sturtevant, 1913a)

Sturtevant (1913a) explique ces manquements à l'additivité pour les grandes distances par l'hypothèse d'une double cassure des chromosomes et par conséquent d'un *double crossing-over* qui annulerait la recombinaison, comme il apparaît clairement dans le schéma de la figure 59.

[...] si une cassure se produit entre B et P , et une autre entre P et M , alors, à moins que l'on puisse suivre P , il n'y aura aucune preuve visible de crossing-over entre B et M , et la mouche qui éclôt du gamète qui en résulte sera rangée dans la classe des absences de crossing-over, alors qu'en réalité elle représente deux *crossing-over*¹⁸. (Sturtevant, 1913a, p. 8)

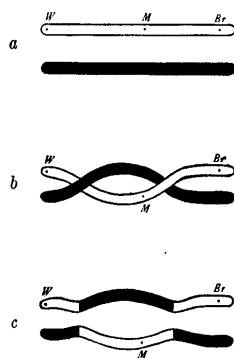


FIG. 59 – Représentation schématique de l'hypothèse du double crossing-over par Morgan *et al.* (1915). Les facteurs W et Br sont hérités ensemble : aucun crossing-over n'est observable, alors qu'il y en a en réalité deux.

En utilisant trois paires ou plus de facteurs liés au sexe dans la même expérience, on peut déterminer si un double crossing-over s'est produit. Sturtevant reproduit dans son article les résultats obtenus par Morgan (1911a) et Morgan et Cattell (1912) lors de telles expériences (voir figure 60).

¹⁸Je souligne.

No crossing over	Single crossing over	Double crossing over	
<div style="display: flex; flex-direction: column; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">B</div> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">CO</div> <div style="border: 1px solid black; padding: 2px;">R</div> </div>	<div style="display: flex; flex-direction: column; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">B</div> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">CO</div> <div style="border: 1px solid black; padding: 2px;">R</div> </div>	<div style="display: flex; flex-direction: column; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">B</div> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">CO</div> <div style="border: 1px solid black; padding: 2px;">R</div> </div>	<div style="display: flex; flex-direction: column; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">B</div> <div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">CO</div> <div style="border: 1px solid black; padding: 2px;">R</div> </div>
6972	3454	60	9

FIG. 60 – Résultats d'expériences faites par Morgan (1911a) et Morgan et Cattell (1912) indiquant des crossing-over doubles pour les facteurs B, CO et R du chromosome X de *Drosophila*. Tableau tiré de Sturtevant (1913a).

En plus de l'hypothèse du double (et parfois triple) crossing-over, Sturtevant observe une autre complication dans l'interprétation des données numériques concernant les recombinaisons observées : il semble que l'occurrence d'un crossing-over en un point empêche ou diminue la probabilité de l'occurrence d'un autre crossing-over en un point proche. C'est ce que Muller (1916) va appeler l'*effet d'interférence* et dont il va tenter de donner une mesure précise par l'élaboration de la notion de « coïncidence ».¹⁹

Les notions de crossing-over multiples et d'interférence vont faire l'objet de travaux développés principalement par Muller (1916). En 1913, Sturtevant n'en donne pas d'explication plus développée. Avant de passer à l'étude de la suite des travaux du groupe de la drosophile sur la cartographie génétique, la technique élaborée par Sturtevant (1913a) appelle plusieurs remarques importantes pour la suite de mon propos.

2.2 Remarques sur les notions d'observation, de crossing-over et de distance

2.2.1 La notion d'observation

En premier lieu, une clarification s'impose sur la notion d'*observation*. La technique de la cartographie génétique repose d'abord sur des données numériques, obtenues par expériences de croisement sur des individus drosophile. Ces données concernent les phénotypes des individus, c'est-à-dire leurs caractères observables. Sans encore prendre en considération les difficultés soulevées par la possibilité de crossing-over multiples, insistons sur un point : les proportions de crossing-over indiquées dans les tableaux sont le fruit d'une première inférence, à partir de l'observation de la distribution des phénotypes chez les hybrides.

Ces données concernant les recombinaisons de facteurs au niveau génotypique,

¹⁹ J'y reviens brièvement plus loin.

dont on dit que ce sont les recombinaisons *observées* (par opposition justement à la possibilité de crossing-over multiples, passant par conséquent inaperçus) ne sont pas elles-mêmes le résultat de l'observation d'un échange de segments au niveau des chromosomes (observation impossible à l'époque), mais sont inférées de la considération des phénotypes.

Ensuite, une seconde étape du raisonnement consiste à convertir des données numériques en indices de distance. Ici intervient la difficulté posée par les crossing-over multiples (et autres manquements à l'additivité, comme les phénomènes d'interférence). En construisant sa carte par la considération des petites distances, Sturtevant obtient, pour les grandes distances, des données concernant les crossing-over *réels* (voir la phrase que j'ai soulignée dans la citation de la page 494), et non les crossing-over observés, au sens défini ci-dessus, c'est-à-dire effectifs (résultant en une recombinaison effective de facteurs, détectable au niveau phénotypique).

Il faut donc bien distinguer entre plusieurs niveaux de phénomènes : 1. le niveau phénotypique, effectivement observé, 2. le niveau génotypique, celui des recombinaisons effectives de facteurs, inférées de 1., 3. le niveau physique du groupe de liaison (qu'il s'agisse ou non du chromosome, comme on va le voir tout de suite), où deux crossing-over réels résultent en une absence de recombinaison effective et n'affectent pas le niveau génotypique.

Or, si c'est bien le niveau génotypique, celui des recombinaisons effectives de facteurs, qui intéresse le généticien, puisque son but est en principe de prédire et d'expliquer la distribution des phénotypes, les distances de la carte représentent la fréquence des crossing-over *réels*, c'est-à-dire au niveau physique (c'est le troisième point, auquel je reviens ci-dessous). Ces remarques sont d'une très grande importance pour la compréhension de la signification des cartes et des débats qu'elles engendrent parmi les généticiens, auxquels je consacrerai une analyse dans les sections 5 et 6.

2.2.2 Le terme de crossing-over

Le deuxième point que je souhaite souligner est une conséquence du premier ; il concerne l'usage du terme même de « *crossing-over* ». Proposé pour la première fois par Morgan et Cattell (1912), il souffre d'une ambiguïté significative. L'échange (hypothétique) entre portions de chromosomes que Morgan (1911c) propose comme explication mécanique de la recombinaison de facteurs observée au niveau génétique (c'est-à-dire par des techniques numériques appliquées aux expériences mendéliennes de croisement) est initialement appelé « chiasmatypie » par Janssens.

Le terme « crossing-over » n'est pas synonyme de « chiasmatypie » : il désigne l'échange entre des portions d'un groupe de liaison, c'est-à-dire d'un ensemble de

gènes liés et ordonnés linéairement, quel que soit le support physique d'un tel groupe de liaison. Pour Morgan, c'est le chromosome ; cependant, il fait clairement la distinction entre la notion de crossing-over et la réalité matérielle d'un échange de portions de chromosomes, comme en témoigne, dans son énoncé de la théorie du gène (voir chapitre 8, page 444), sa mention du crossing-over qu'il définit comme « un échange ordonné [...] entre les éléments de groupes de liaison correspondants » (Morgan, 1928), qui n'implique pas, en tant que telle, d'engagement en faveur de la théorie chromosomique.

De plus, une autre ambiguïté encore plus importante affecte l'usage du terme « crossing-over » : il est en effet quelquefois employé pour désigner la recombinaison de facteurs elle-même (on parle alors de « crossing-over *observé* ou *effectif* ») en tant qu'elle peut être déduite des observations au niveau phénotypique (elle est « observable » au sens défini à la section précédente). Dans ce cas, deux crossing-over au sens d'échanges de portions d'un groupe de liaison (crossing-over *réels*) comptent, en ce qui concerne les gènes situés aux extrémités, pour une absence de crossing-over au sens de recombinaison de facteurs. Le terme, par sa connotation mécanique, semble bien désigner ce qui se passe au niveau physique (chromosomique ou non) ; cependant, dans les débats engendrés par les travaux de l'école de Morgan, dont certains remettent en cause la base physique (et pas seulement chromosomique) de la théorie du crossing-over, le terme est employé par des opposants (c'est ce qui permettra, par exemple, à Goldschmidt (1917) de plaider pour un « crossing-over sans chiasmatisation », c'est-à-dire un échange de facteurs sans échange de portions de leur support matériel) ; selon qu'ils acceptent ou non la possibilité d'un double échange entre morceaux d'un groupe de liaison – et non pas nécessairement d'un chromosome –, le terme de crossing-over pourra désigner les échanges physiques qui, en nombre pair, annulent leurs effets génotypiques, ou les recombinaisons effectives, quelles que soient leurs causes mécaniques (j'y reviendrai plus loin).

2.2.3 La notion de distance

Le troisième point sur lequel je souhaite insister découle des deux précédents : la notion de *distance* est définie (voir page 492) comme le pourcentage de crossing-over. Étant donné la possibilité de crossing-over multiples, le pourcentage à prendre en compte est celui des crossing-over physiques ; c'est pourquoi les prédictions obtenues pour les grandes distances ne sont pas les bonnes, et les distances sont établies sur la base des petites distances, évitant ainsi les effets des crossing-over multiples.

Ainsi, la distance n'est pas simplement la probabilité de recombinaison, mais celle d'une cassure « réelle », quelle que soit sa nature physique et quelles qu'en soient les

conséquences au niveau génétique. La distance n'est par conséquent pas une fonction linéaire du nombre de recombinaisons *effectives*, mais du nombre de crossing-over *réels*. En ce sens, avant même la confirmation de l'hypothèse chromosomique, la notion de distance ne correspond pas à la fréquence de recombinaison de facteurs génétiques, mais a bien la signification d'une distance physique.

Malgré cela, et c'est là une quatrième difficulté, la notion de distance ainsi définie n'est pas censée correspondre précisément à une distance physique réelle. Dès 1913, en effet, Sturtevant est très conscient du fait que le concept de distance génétique ou cartographique ne correspond pas nécessairement à une distance physique sur le chromosome :

Bien entendu, on ne sait pas si ces distances ainsi obtenues représentent la distance relative réelle qui sépare les facteurs. Ainsi, la distance CP [comprendre : la distance physique] peut en réalité être plus courte que la distance BC , mais ce que l'on sait est qu'une cassure est beaucoup plus probable entre C et P qu'entre B et C . Donc, soit CP est une portion [physique] longue, ou alors, pour quelque raison, c'est une portion fragile. Ce que je souhaite dire ici est que nous n'avons pas de moyen de savoir si les chromosomes sont uniformément solides, ou s'ils ont des portions plus ou moins fragiles, ce qui empêcherait notre diagramme de représenter des distances relatives réelles – sans pour autant, je pense, que cela lui ôte sa valeur *en tant que diagramme*²⁰. (Sturtevant, 1913a, p. 6)

Ainsi, distance génétique et distance physique, qu'elles aient ou non la même valeur, ne sont pas des concepts de même nature : la première est un concept probabiliste, la seconde le résultat d'une mesure physique. Cette précaution de Sturtevant sera d'ailleurs confirmée par les avancées en cytologie : les chromosomes sont par exemple beaucoup moins susceptibles de se casser au niveau du centromère ; la distance est donc une mesure de solidité autant que de longueur.

Insistons sur l'avancée théorique cruciale que constitue cette définition de la distance en termes de pourcentage de crossing-over. Elle est opérée en cours d'article par Sturtevant. En effet, au début de l'article, après avoir exposé l'hypothèse de Morgan, pour introduire la technique qu'il y met en œuvre, Sturtevant suggérait de prendre le nombre de crossing-over comme un *indice* de la distance et fondait l'additivité des proportions de crossing-over sur l'hypothèse selon laquelle ces dernières « représentent réellement la distance » (voir page 491). Dans cette dernière occurrence, le terme de distance semble désigner la distance physique réelle entre deux loci du chromosome. Après avoir défini la distance comme le pourcentage de crossing-over, une telle modalisation n'a plus lieu d'être. D'une hypothèse au niveau cytologique, des-

²⁰ Je souligne.

tinée à fournir une explication mécanique d'un phénomène génétique, Sturtevant a tiré un nouvel ensemble de concepts génétiques indépendants du niveau cytologique.

Je reviendrai largement sur ces remarques dans ce qui suit ; je voulais cependant dès à présent insister sur la très grande complexité et quelquefois l'ambiguïté des termes de « distance » et de « crossing-over », car cette ambiguïté est la conséquence et le signe des différents niveaux d'analyse en jeu et de la grande abstraction de l'analyse génétique.

3 Le travail du groupe de la drosophile

Les années qui suivent sont celles des avancées les plus importantes et parmi les plus célèbres de l'histoire de la génétique classique : le « groupe de la drosophile », dans le laboratoire de Morgan à Columbia, en collaboration étroite avec le laboratoire du cytologiste Wilson, construit progressivement la carte génétique des quatre chromosomes de la drosophile (figure 62, page 502). Au moyen de prédictions remarquables au niveau génétique, les élèves de Morgan apportent des preuves de plus en plus importantes en faveur de la thèse du support chromosomique des gènes, qui évince progressivement la théorie concurrente de Bateson.²¹ Le *Mechanism of Mendelian Heredity* de 1915 se présente comme une défense de la théorie chromosomique de l'hérédité, hypothèse mécanique expliquant les données mendéliennes. Les thèses de doctorat de Calvin Bridges (1916a,b), Hermann J. Muller (1916) et Alfred Sturtevant (1913a, 1914, 1915) contiennent chacune des résultats cruciaux qui constituent les avancées majeures de la génétique classique. Je n'en rapporterai que les points indispensables à la compréhension de mon propos.²²

3.1 La carte génétique complète de la drosophile

La carte de Sturtevant (1913a) est celle du chromosome *X*. C'est la plus facile à construire, en raison de l'absence d'un second chromosome *X* chez le mâle. En effet, dans la mesure où le seul chromosome *X* d'un mâle lui vient de sa mère, les caractères liés au sexe des mâles d'une génération indiquent sans complication le type d'œufs produits par les femelles de la génération antérieure (voir Morgan et Cattell, 1912). Pour construire la carte des autres chromosomes, Morgan et ses élèves utilisent les fréquences de crossing-over dans le cas des facteurs liés au chromosome *X* comme standard pour déterminer les fréquences de crossing-over des facteurs non liés au

²¹Voir (Carlson, 1967, chap. 7).

²²Il existe de très nombreuses histoires de cette période, dont certaines écrites par ses principaux acteurs ; voir (Dunn, 1965 ; Sturtevant, 1965 ; Carlson, 1967 ; Allen, 1978 ; Mayr, 1982 ; Kohler, 1994).

sexe. La technique clé de Morgan et de son groupe est celle du *back-cross*, illustrée par la figure 61.²³

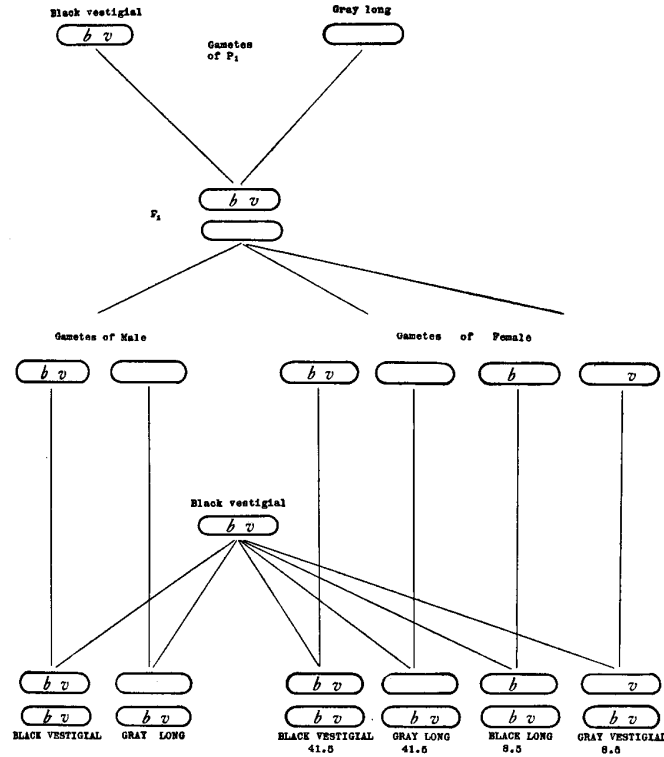


FIG. 61 – La technique du *back-cross*. Le test *back-cross* illustré ici vise à mesurer la distance entre le gène responsable de la couleur noire du corps (*b* pour « *black* ») et le gène responsable des ailes atrophiées (*v*, pour « *vestigial* »). On produit d'abord un stock de mouches homozygotes doublement récessives (*bv bv*). Ce sont ces mouches qui permettront de rendre les crossing-over observables. On croise alors ce stock avec des mouches de type sauvage. Dans le diagramme, l'absence de lettre (le blanc) représente les gènes sauvages; pour plus de clarté, je les note + dans ce qui suit. C'est au sein de la première génération *F*₁ d'hybride femelles hétérozygotes (*bv ++*) qu'un crossing-over va se produire pendant la gamétogénèse, produisant des œufs *b+* et *+v*. Pour que ces recombinaisons soient visibles, on croise en retour (*backcross*) les femelles *F*₁ avec des mâles (*bv bv*). Au sein de la progéniture de ce croisement (*F*₂), certaines mouches seront noires ou auront les ailes atrophiées mais ne présenteront pas les deux caractères. Ces mouches, (*b+* *bv*) et (*+v* *bv*), sont celles chez qui il y a eu recombinaison (Morgan *et al.*, 1915, pp. 49-51).

La cartographie des trois autres chromosomes de la drosophile a été rendue possible par l'identification progressive de trois autres groupes de liaison entre facteurs. En 1912, Morgan et Lynch avaient déjà trouvé un groupe de mutants non liés au sexe :

²³Voir (Kohler, 1994, p. 70-71) et (Carlson, 1967, p. 54) pour plus de détails sur les techniques mises en œuvre par le groupe de Morgan.

celui des facteurs responsables de la couleur noire du corps et des ailes atrophiées. Ce sont ceux qui se trouvent sur le deuxième chromosome. Sturtevant (1913b) montre l'existence d'un troisième groupe de gènes, qui permet d'établir la carte du troisième chromosome. Enfin, Muller (1914) trouve « un gène pour le quatrième chromosome de *Drosophila* ». Les observations de Stevens et Wilson avaient montré que le quatrième chromosome de la drosophile était de très petite taille. Or les facteurs non liés aux trois autres groupes de liaison apparaissent complètement liés entre eux, ce qui semble concorder avec la petite taille du quatrième chromosome. La correspondance des quatre groupes de liaison avec les quatre paires de chromosomes dont les observations de Stevens et Wilson ont rapporté l'existence chez la drosophile apporte un argument de poids en faveur de l'hypothèse chromosomique. Le nombre de gènes sur chaque chromosome, leur distance cartographique, et la remarquable correspondance entre le chromosome *X* et les facteurs liés au sexe sont autant d'arguments en faveur de cette hypothèse. Muller (1914) peut ainsi écrire glorieusement que « le fossé principal qui demeurerait dans la série des phénomènes génétiques traçant un parallèle avec les faits cytologiques connus chez *Drosophila* a [...] été comblé. »

3.2 Vers la confirmation de l'hypothèse chromosomique

Comme on vient de le voir, la cartographie des quatre chromosomes de *Drosophila* par le groupe de Morgan s'accompagne de la défense de la théorie chromosomique de l'hérédité. Les thèses de doctorat de chacun des élèves de Morgan, dans les années 1915 et 1916, constituent des avancées considérables en direction de la confirmation de cette théorie. Je me contenterai, ici, d'en mentionner rapidement les épisodes les plus significatifs.²⁴

Au moment où Sturtevant construit la première carte génétique, les travaux de Bridges (1913) sur la non-disjonction apportent les premières preuves cytologiques importantes en faveur de l'hypothèse chromosomique. En effet, Calvin Bridges (1913), constatant des anomalies au niveau *génétique*²⁵ les interprète comme le résultat d'une

²⁴Pour toute cette période, voir (Sturtevant, 1965, Dunn, 1965, Carlson, 1967, chap. 9, Darden, 1991, Falk, 2004).

²⁵Ces anomalies se produisent dans certains croisements de type « criss-cross ». Ce type de croisements, baptisés ainsi parce que les mâles sont identiques à leur mère et les femelles à leur père, consistent en un croisement entre des femelles aux yeux blancs et des mâles aux yeux rouges, et donnent les résultats suivants : à la génération F_1 , les femelles ont les yeux rouges et les mâles ont les yeux blancs. Dans certains cas, Bridges rapporte que 5% des femelles sont identiques à leur mère et 5% des mâles à leur père. Les croisements critiques impliquent une femelle de la génération parentale P_1 qui a un facteur sur le troisième chromosome, « yeux roses », et des mâles aux yeux blancs et ailes miniatures sur le chromosome *X* et le corps de couleur noir (deuxième chromosome). Voir (Carlson, 1967, p. 68).

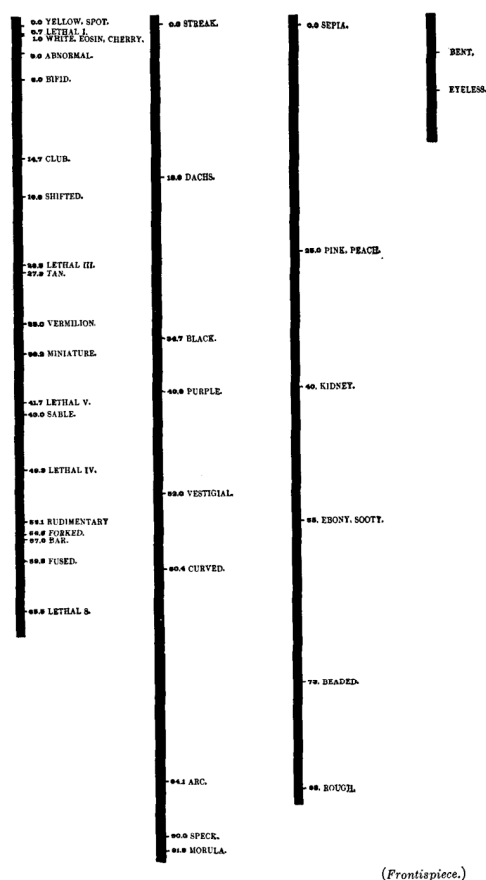


FIG. 62 – Carte génétique des quatre chromosomes de la drosophile (Morgan *et al.*, 1915)

aberration au niveau chromosomique, qu'il appelle « non-disjonction » : dans de rares cas, les chromosomes *X* de la femelle ne se séparent pas pendant la méiose.²⁶ L'observation au microscope de ces phénomènes de non-disjonction qui expliquent les anomalies constatées au niveau génétique, en confirmant l'hypothèse de Bridges, lui permet d'affirmer, en 1916, avoir fourni une « preuve directe, par la non-disjonction, que les gènes liés au sexe de *Drosophila* sont portés par le chromosome *X*. »

D'autre part, Muller (1916) analyse en détail les mécanismes de recombinaison et d'interférence. La pièce centrale en est l'étude des recombinaisons de douze facteurs

²⁶ Dans de tels cas, les deux *X* de la mère se retrouvent alors dans une même cellule et l'autre cellule ne comporte pas de *X*. Un œuf portant deux chromosomes *X*, rencontrant le sperme déterminant pour un mâle (qui ne contient pas de *X*), produit une femelle. Un œuf sans chromosome *X*, fertilisé par du sperme déterminant une femelle (portant le chromosome *X* du père), produit un mâle qui, en violation de l'hérédité criss-cross (voir note 25), est identique à son père (pour la couleur des yeux).

mutants sur le chromosome X . Au moyen d'un très grand nombre d'expériences de croisement et de procédures statistiques, il cherche à mesurer le plus précisément possible l'effet d'interférence, constaté dès 1913 par Sturtevant, en détectant avec précision tous les crossing-over simples et multiples. Il développe le concept de coïncidence, comme mesure de l'interférence, et l'ensemble des procédures pour la calculer. La coïncidence est le ratio entre les crossing-over observés et les crossing-over attendus calculés sur la base des crossing-over observés précédemment (figure 63). Dans les procédures mises en œuvre par Muller, la signification des notions de « distance », d'« observation » et de « crossing-over » gagne encore en complexité, et mériterait un traitement détaillé. Je me contente de signaler ces travaux, et renvoie à l'article de William Wimsatt (1992) pour le détail de la procédure de Muller.²⁷

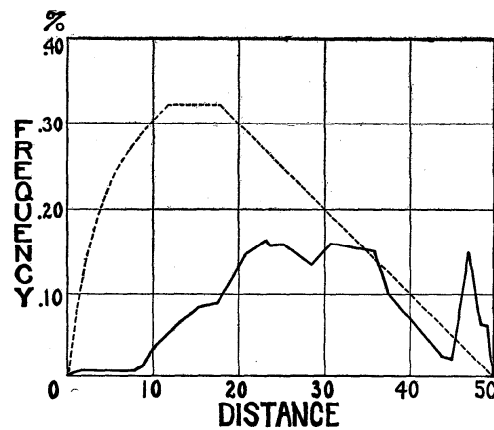


FIG. 63 – Graphe des fréquences attendues et « observées » de doubles crossing-over en fonction de la distance pour le chromosome X . (Muller, 1916, p. 424)

Malgré ces remarquables avancées du groupe de la drosophile en direction de la confirmation de l'hypothèse chromosomique, qui gagne progressivement du terrain contre l'hypothèse de la reduplication de Bateson, il convient d'insister sur le fait que, jusqu'au début des années 1930, *aucune preuve cytologique d'un échange effectif* entre portions de chromosomes au moment de la méiose n'est disponible. La non-disjonction observée par Bridges consiste en effet en l'absence de séparation du chromosome X , et ne constitue pas un cas de chiasmotypie. Si le regard historique rétrospectif sur les travaux du groupe de Morgan encourage à y voir des preuves très convaincantes en faveur de l'hypothèse chromosomique, en raison des analogies frap-

²⁷Wimsatt (1992) analyse la méthode de Muller en regard de celle de Haldane qui, en 1919, propose une fonction mathématique de la distance cartographique, idée à laquelle le groupe de Morgan reste réticent.

pantes entre le comportement des chromosomes et celui des gènes, il convient de ne pas négliger le fait que l'idée même selon laquelle les gènes sont des entités matérielles est loin d'aller de soi pour de nombreux généticiens.²⁸ On peut par conséquent reconnaître les résultats numériques du groupe de Morgan, voire admettre l'hypothèse du crossing-over, c'est-à-dire l'hypothèse d'un échange entre *facteurs héréditaires*, sans pour autant considérer que les gènes sont des parties de chromosomes. C'est le cas, en particulier de William Bateson, généticien de la première heure, qui reste un opposant de l'hypothèse chromosomique jusqu'en 1921.

Le modèle du collier de perles (figure 55), tout en prenant la forme d'un chromosome (et donc en se présentant comme un schéma de chromosome – voir chapitre 8, page 454), n'est pas censé représenter la structure réelle des chromosomes. Les gènes y sont représentés par analogie avec les perles d'un collier, et rien n'est dit de la nature de ces « perles » ni des modalités de leur localisation sur les chromosomes. La valeur explicative de ce modèle, tout en étant due à son aspect schématique, est encore hypothétique.

Je reviendrai, aux sections 5 et 6, à la question du rôle de cette hypothèse mécanique, et de la manière dont elle joue sur la compréhension que les différents généticiens des années 1920 ont des phénomènes de crossing-over et sur leur utilisation et leur interprétation des cartes génétiques. Auparavant, je voudrais rapidement présenter les travaux des années 1930 qui ont apporté des preuves cytologiques en faveur de la théorie chromosomique suffisantes pour en convaincre presque tous ceux qui s'y opposent encore dans les années 1920.²⁹

4 Les années 1930 : confirmation de l'hypothèse chromosomique

Les travaux de Bridges (1913, 1916a,b) sur la non disjonction ont apporté de fortes preuves en faveur de l'hypothèse chromosomique, en établissant une corrélation entre anomalies génétiques et aberrations cytologiques. Cependant, comme je l'ai dit, on ne dispose, à la fin des années 1920, d'aucune preuve de la correspondance entre recombinaisons de facteurs et *échanges de portions de chromosomes*, dans la mesure

²⁸De plus, au début des années 1910, rien ne laisse encore soupçonner que la structure fine des chromosomes est effectivement linéaire. C'est au milieu des années 1910 que des découvertes cytologiques suggèrent une forme linéaire, contre l'hypothèse selon laquelle les chromosomes ont la forme d'une tige pendant la métaphase et l'anaphase et se fragmentent à la fin de la division et se dissolvent en un réseau de chromatine qui se réagrège pendant la division suivante.

²⁹À l'exception, en particulier, de Goldschmidt, qui en reste un opposant jusque dans les années 1950.

où l'on ne dispose pas encore de preuve cytologique d'un échange effectif de morceaux entre les chromosomes lors de leurs entrecroisements au cours de la méiose. Cette preuve est apportée simultanément en Curt Stern (1931) pour la drosophile et par Harriet B. Creighton et Barbara McClintock (1931) pour le maïs.³⁰ Leurs travaux montrent de plus que ces échanges de portions de chromosomes correspondent, au niveau génétique, à des recombinaisons de facteurs. Mais c'est seulement en 1934 que la correspondance entre certains facteurs et des portions précises de chromosomes peut être visuellement établie, grâce à l'observation des chromosomes géants des glandes salivaires de la drosophile par Painter (1934a).

4.1 Creighton et McClintock (1931) : crossing-over cytologiques et crossing-over génétiques

Les expériences menées par Creighton et McClintock (1931) sur *Zea mays* leur permettent d'administrer une double preuve : premièrement, elles montrent qu'un échange entre portions de chromosomes homologues – qu'elles appellent « *crossing-over cytologique* » – a effectivement lieu pendant la méiose ; deuxièmement, elles montrent que cet échange est accompagné d'une recombinaison de facteurs au niveau génétique – qu'elles appellent « *crossing-over génétique* ».

Pour ce faire, il faut parvenir à isoler des paires de chromosomes homologues hétéromorphes, c'est-à-dire présentant des différences physiques visuellement détectables. Ces différences jouent le rôle de marqueurs cytologiques permettant de repérer d'éventuels échanges, de même que les différents allèles d'un même gène sont des marqueurs génétiques :

L'étude génétique des crossing-over exige une condition hétérozygote de deux facteurs alléomorphiques dans le même groupe de liaison. L'analyse du comportement de chromosomes homologues ou partiellement homologues, qui sont morphologiquement distinguables en deux points, doit nous fournir la preuve d'un crossing-over cytologique. (Creighton et McClintock, 1931, p. 1)

Creighton et McClintock obtiennent ainsi, par rayonnement, une espèce de maïs dont un des chromosomes de la paire n°9 est anormal : on repère une boursouffure à l'une de ses extrémités et des translocations³¹ sont détectables à l'autre extrémité en raison d'une inégalité de taille des morceaux échangés (voir figure 64). Par des expériences de croisement et des observations à chaque génération, elles parviennent à s'assurer que la distribution de la boursouffure au sein des générations successives est

³⁰Pour une analyse des travaux de Barbara McClintock, voir (Kass et Bonneuil, 2004).

³¹Il s'agit d'un échange entre le chromosome n°8 et le chromosome n°9, repéré par McClintock (1930).

similaire à celle d'un gène et que « la boursouffure est une caractéristique constante du chromosome qui la possède ». Croisant des individus pourvus du chromosome anormal avec des individus normaux, elles obtiennent des individus ayant des paires de chromosomes hétéromorphes.

Quand elle est présente sur un chromosome et pas sur son homologue, la boursouffure rend la paire de chromosome visiblement hétéromorphe. (Creighton et McClintock, 1931, p. 2)

Elles parviennent ainsi à montrer qu'un crossing-over cytologique se produit entre la boursouffure et la translocation dans environ 39% des cas.

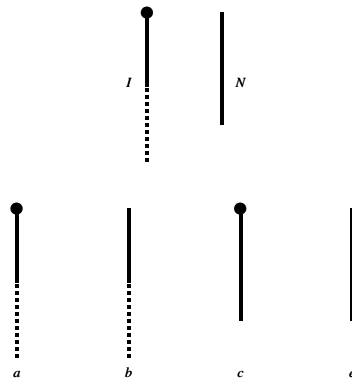


FIG. 64 – Schéma des chromosomes étudiés par Creighton et McClintock (1931). *I*, chromosome ayant subi l'échange; *N*, chromosome normal. *a*, chromosome ayant subi l'échange, boursoufflé; *b*, chromosome ayant subi l'échange, sans boursouffure; *c*, chromosome normal, boursoufflé; *d*, chromosome normal, sans boursouffure. *a* et *d* sont des types sans crossing-over. *b* et *c* sont des types avec crossing-over.

Le deuxième temps de leur article est consacré à montrer que ces crossing-over cytologiques correspondent à des recombinaisons de gènes connus pour être portés par ce chromosome : les gènes *C* (aleurone coloré, *colored aleurone*), *sh* (endosperme réduit, *shrunk endosperm*) et *wx* (endosperme cireux, *waxy endosperm*). McClintock (1930) avait en outre montré que ces gènes sont alignés à partir du point de translocation, dans l'ordre suivant : $wx - sh - c$. Elle avait alors établi une correspondance entre les emplacements relatifs des gènes et les portions du chromosome de part et d'autre du point de translocation (voir figure 65).

Creighton et McClintock procèdent aux croisements suivants :

$$K - C - wx - I - / k - c - Wx - i \times k - c - Wx - i / k - c - wx - i.$$

La représentation diagrammatique de leur expérience est donnée dans la figure 66 (pour le détail du protocole expérimental, voir Creighton et McClintock, 1931, p. 4-5). Leurs résultats sont présentés numériquement dans le tableau de la figure 67.

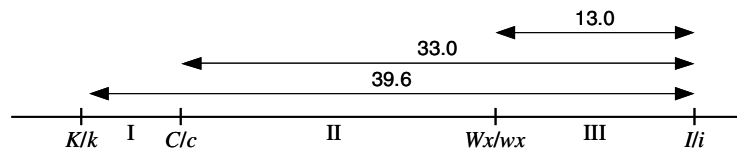


FIG. 65 – Carte cytologique et génétique (McClintock, 1930).



FIG. 66 – Représentation diagrammatique du croisement (Creighton et McClintock, 1931, p. 5).

Knob-C-wx Knobless-c-Wx				x	Knobless-c-wx			
C-wx		c-Wx			C-Wx		c-wx	
Knob	Knobless	Knob	Knobless		Knob	Knobless	Knob	Knobless
12	5	5	34		4	0	0	3

FIG. 67 – Résultats numériques (Creighton et McClintock, 1931, p. 5)

Ces résultats permettent de montrer que « des paires de chromosomes, hétéromorphes en deux régions, échangent des portions en même temps qu'ils échangent des gènes assignés à ces régions. » (Creighton et McClintock, 1931, p. 6) Autrement dit, ils apportent la preuve de la correspondance entre crossing-over cytologiques (échanges effectifs de portions de chromosomes) et crossing-over génétiques (recombinaisons).

4.2 Painter (1934a,b) : la correspondance des cartes génétiques et des cartes cytologiques

En 1934, Theophilus Painter met au point une nouvelle technique d'observation cytologique pour étudier la morphologie des chromosomes géants des glandes salivaires de la drosophile. Elle lui permet de détecter et de mesurer avec précision les différentes aberrations chromosomiques obtenues par irradiation : translocations, délétions et inversions (voir figure 70).

Les éléments ont une morphologie si distincte et constante que l'on peut suivre un segment qui subit une translocation jusqu'à son nouvel emplacement, ou, si l'on étudie une translocation mutuelle, on peut déterminer précisément la taille

de la portion de chaque chromosome concernée par l'échange. (Painter, 1934b, p 175)

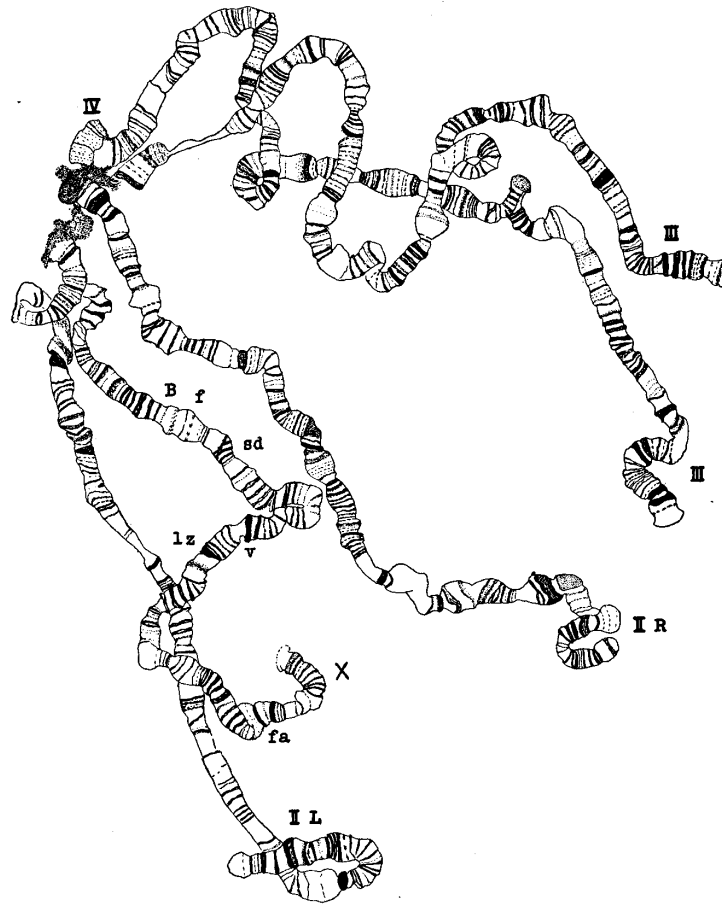


FIG. 68 – Image de chromosome de la grande salivaire de *Drosophila* obtenue à la chambre claire (*camera lucida*), à partir d'une préparation acéto-carmin. II – L et II – R indiquent respectivement les brins de gauche et de droite du deuxième chromosome. (Painter, 1934b, p .179)

La particularité de ces chromosomes est que les homologues s'assortissent par paires rangées, ligne par ligne, bande par bande, et s'unissent en un seul élément. Par conséquent, si un des chromosomes subit une inversion, il devient possible d'en déterminer l'emplacement exact. La figure 68 présente ainsi un dessin tiré d'une image obtenue à la chambre claire grâce à la technique de Painter (1934b) pour des chromosomes sans aberration. Les différentes images de la figure 70 exposent la manière dont sont cartographiées les différentes portions de chromosome, sur la base des aberrations visibles grâce aux alternances de bandes colorées (pour un exposé détaillé du protocole, voir Painter, 1934b).

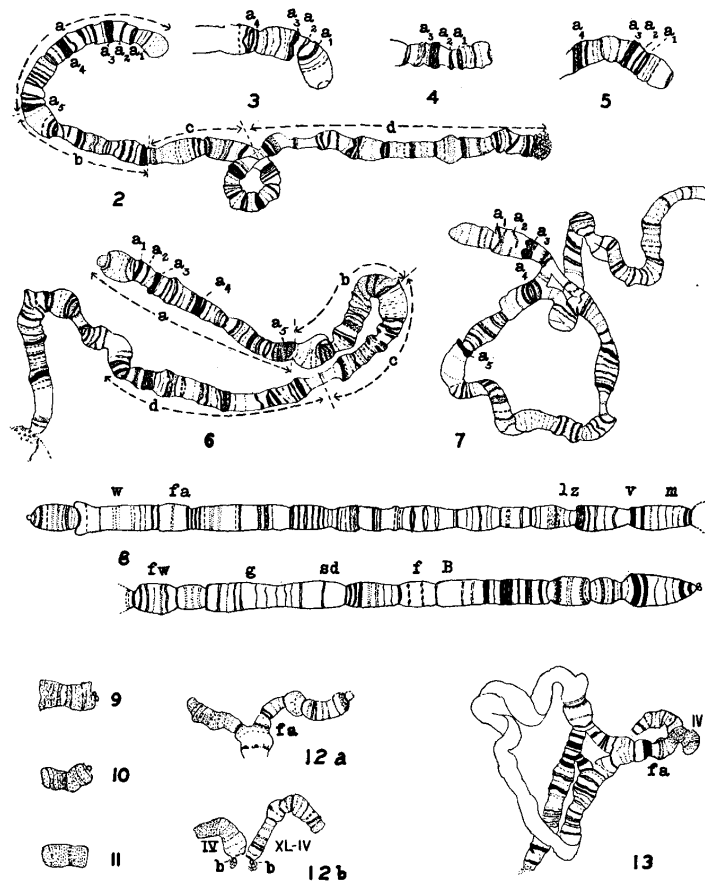


FIG. 69 — Les figures 2 et 6 présentent la morphologie typique du bras gauche du deuxième chromosome. Les figures 3 à 5 présentent les segments des extrémités prises dans d'autres cellules. La figure 7 présente l'inversion trouvée dans le bras gauche du deuxième chromosome quand la larve est hétérozygote pour le gène *curly*. La figure 8, obtenue par la combinaison d'images de chambre claire, présente la structure typique du chromosome X. L'emplacement approximatif d'un certain nombre de loci de gènes, tels qu'il est déterminé par l'étude des cassures, est indiqué par les symboles génétiques usuels. Les figures 9 et 11 présentent la morphologie de la paire des quatrièmes chromosomes. Les figures 12a et 12b, prises de la même cellule, présentent une translocation mutuelle entre le chromosome X et le quatrième chromosome. La figure 13 est l'inversion obtenue, quand une larve est hétérozygote pour ClB est examinée. Dans ce cas, le chromosome X normal porte un morceau du quatrième chromosome sur son extrémité de gauche. (Painter, 1934b, p. 182)

Dans un deuxième article, Painter (1934a) applique cette technique au chromosome X de la drosophile, et fait correspondre les différents loci génétiques aux portions de chromosomes identifiés visuellement. Cela lui permet d'établir un « nouveau type de carte chromosomique [...] montrant la position approximative de 21 loci de gènes disposés le long de la région active du X. » (Painter, 1934a, p. 448)

Ce nouveau type de carte consiste en une carte chromosomique (*i.e.* obtenue par sa technique d'observation cytologique), à laquelle il fait correspondre la carte génétique proprement dite. Il poursuit la même opération pour les autres chromosomes. La figure 70 représente ainsi la correspondance des cartes génétique et cytologique du troisième chromosome.

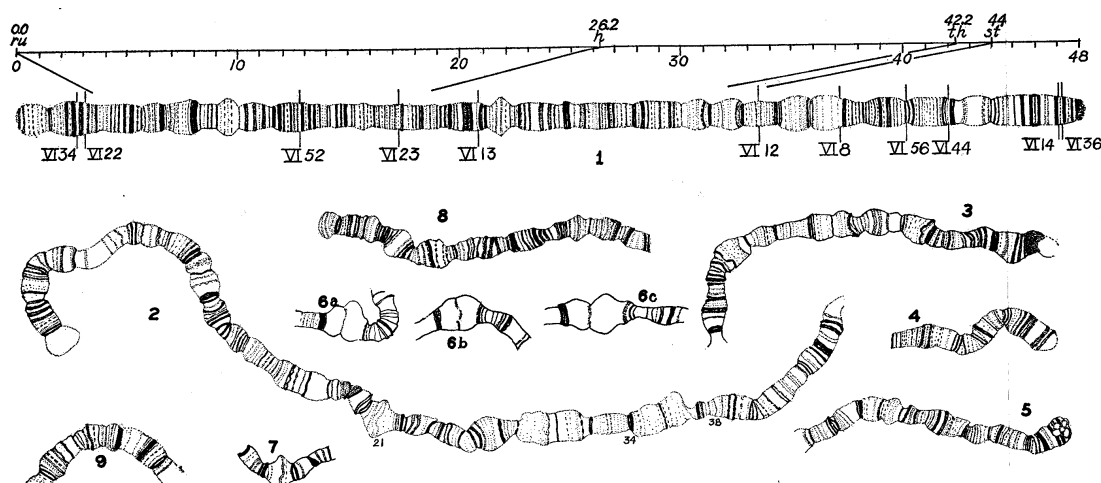


FIG. 70 – Carte cyto-génétique du troisième chromosome de la drosophile. 1. Carte cytologique de la morphologie typique de la portion euchromatique du bras gauche du troisième chromosome. Le graphe au-dessus est une carte de crossing-over (génétique). Des lignes relient les gènes identifiés par croisement aux limites des emplacements des loci sur les chromosomes correspondants. En dessous, les nombres se réfèrent aux différentes translocations étudiées. 2. Image de la portion euchromatique du bras gauche intact. Les chiffres placés le long du dessin aident le lecteur à identifier les divers endroits correspondants sur la carte cytologique de la figure 1. 3 à 5. Topographie des fibres des extrémités de la zone euchromatique. 6a à 6c. Variations de la taille des élargissements achromatiques entre 34 et 37 sur l'échelle cytologique. 7. Elargissement en forme de fuseau au niveau du 21 sur l'échelle des crossing-over. 8 et 9. Topographie de l'extrémité de gauche du bras de gauche. (Painter, 1935, planche entre les pages 308 et 309)

Avant de revenir aux débats engendrés par la technique de cartographie de Morgan dans les années 1920, quelques remarques sur la nature de ces représentations s'imposent. La technique de Painter permet, comme on vient de le voir, de faire correspondre les cartes génétiques avec les cartes cytologiques. Les premières, obtenues par des techniques mendéliennes, sont, pour reprendre la terminologie définie au chapitre 7, des *graphes*, c'est-à-dire des sortes de *diagrammes* : les relations spatiales représentent des données numériques. En tant que telles, elles permettent de visualiser ces données, mais ne contiennent pas plus d'informations qu'elles. Elles permettent de représenter graphiquement ces données numériques afin de faciliter certaines inférences. Ce ne sont donc des « cartes » qu'en un sens métaphorique :

elles ne suffisent pas à conclure que les distances génétiques correspondent effectivement à des distances physiques, ni que les relations spatiales de la carte correspondent à des relations spatiales. Les cartes cytologiques, elles, sont obtenues par des techniques d'imagerie, et appartiennent donc à la catégorie des *schémas* (représentations maintenant des relations topologiques).

Malgré le caractère approximatif de l'assignation des gènes aux loci, la mise en correspondance des deux cartes peut être considérée comme attribuant à la carte génétique un statut différent. De carte diagrammatique (de graphe) elle devient une carte schématique, c'est-à-dire une représentation, quoique imprécise, de l'emplacement des gènes sur les chromosomes (une carte en un sens non métaphorique). Jusqu'aux expériences de Painter, voir dans les cartes génétiques des cartes chromosomiques est un acte de foi. En 1934, la correspondance entre les deux cartes est établie. Je me propose à présent, par un « retour en arrière » d'étudier la manière dont les engagements des différents généticiens dans les années 1920 jouent dans leur lecture et leur interprétation des cartes. On verra notamment que le modèle analogique du collier de perles et l'hypothèse chromosomique jouent un rôle, pour Morgan et ses élèves, que l'on peut difficilement restreindre à un rôle « heuristique » ; je vais en effet montrer que la technique de construction des cartes elle-même en dépend.

5 Les critiques de Goldschmidt (1917), Bateson (1916) et Castle (1919)

Les travaux du groupe de Morgan suscitent de nombreuses critiques, entre la parution de (Morgan *et al.*, 1915) et le début des années 1920, date à laquelle, à part quelques irréductibles comme Goldschmidt, tous leurs opposants initiaux, parmi lesquels William Bateson et William Castle, sont convaincus par leur modèle. Auparavant, les critiques les plus importantes leurs sont faites par Bateson, Goldschmidt, Castle et Punnett. Je me propose, dans cette section, d'évoquer rapidement les deux premières, avant de me pencher plus en détail sur la troisième (celle de Castle). Comme on va le voir, chacun de ces généticiens accepte certaines hypothèses du groupe de Morgan et en refuse d'autres (le tableau 2, page 525) récapitule leurs positions respectives. Je tirerai les conclusions de ces débats dans la section 6.

5.1 Goldschmidt (1917) : Crossing-over sans chiasmatypie

Richard Goldschmidt est un des opposants les plus fervents à la génétique du groupe de Morgan. Une des raisons de cette opposition est due au fait que Goldschmidt refuse la distinction méthodologique fondamentale entre l'étude de l'hérédité et

celle du développement³² ; une théorie du gène doit, selon lui, dire quelque chose du mécanisme d'action des gènes. Or, comme on l'a vu, les recherches de Morgan naissent du constat qu'il dresse, au début des années 1910³³, de la nécessité méthodologique de cette distinction.

En conséquence, le modèle qu'il propose est destiné à accorder une place importante à l'action génique.³⁴ Tout en acceptant deux des hypothèses fondamentales de l'école de Morgan – par ailleurs très contestées –, l'hypothèse chromosomique et celle de l'ordonnancement linéaire des gènes sur les chromosomes, Goldschmidt (1917) refuse la thèse d'un échange de portions de chromosomes (chiasmotypie). Il admet donc l'hypothèse d'un crossing-over *génétique*, c'est-à-dire d'un échange de facteurs ou gènes entre les chromosomes, mais refuse de corrélérer ce phénomène à un échange *physique*, appelé « chiasmotypie » depuis Janssens (1909) – ou « crossing-over cytologique » pour reprendre les termes de Creighton et McClintock (1931). Pour expliquer un tel crossing-over génétique sans support physique, il suppose que les chromosomes se dissolvent entre les divisions cellulaires ; à ce moment, les gènes se rendent dans le cytoplasme pour causer leurs effets phénotypiques. Puis, au cours de la division cellulaire suivante, les chromosomes se reconstituent et certains gènes sont déplacés.

L'explication du déplacement relatif des gènes se fait par l'appel à un concept de « force » : moins deux allèles sont identiques du point de vue de leur « force », plus ils sont « proches » sur la carte. La carte génétique est donc une représentation commode des forces relatives des gènes mais n'a aucune relation avec les positions réelles des gènes sur leur support chromosomique. Les facteurs ne sont pas des portions de chromosomes, mais sont attachés à des emplacements spécifiques du chromosome par des forces de nature biochimiques. Pour accomplir leur action au sein de la cellule, les facteurs abandonnent le chromosome. Chaque facteur a une force spécifique, et les différents allèles ont des forces différentes, ce qui semble s'accorder avec l'action génique.

Sturtevant (1917), dans sa réponse à Goldschmidt, souligne le fait que l'existence de crossing-over multiples, mise en évidence au niveau génétique – et que Goldschmidt doit accepter en raison de son adoption de l'hypothèse de linéarité – est un argument contre le modèle de Goldschmidt. Selon ce dernier, en effet, les échanges de facteurs sur un locus devraient être indépendants des échanges de facteurs sur un autre locus. Dans le modèle de Sturtevant, en revanche, l'échange d'un facteur s'accompagne de l'échange des facteurs situés à proximité sur la carte.

³²Pour une étude de la génétique physiologique et développementale de Goldschmidt, voir (Allen, 1974; Dietrich, 2000).

³³Jusque dans les années 1910, il refuse cette distinction.

³⁴Pour une analyse détaillée du modèle de Goldschmidt (1917), voir (Wimsatt, 1987).

Sans détailler le débat entre Goldschmidt et Sturtevant, je voudrais tirer quelques remarques du résumé qui précède. La position de Goldschmidt met en évidence le fait qu'aucune hypothèse constitutive de la théorie de Morgan ne va de soi pour les généticiens de l'époque : l'opposition à Morgan ne prend pas obligatoirement la forme d'un rejet de la théorie chromosomique, mais cette dernière, à l'époque, peut elle-même prendre plusieurs formes. En l'absence d'un accès à la structure fine des chromosomes, l'hypothèse chromosomique elle-même peut signifier diverses choses : les gènes peuvent être considérés comme des unités localisées sur les chromosomes, selon diverses modalités (Goldschmidt suggère qu'ils sont attachés par des forces biochimiques, et l'on pourrait imaginer de nombreuses autres sortes de lien), ou même comme des portions de chromosomes. Aucune connaissance cytologique ne permet, à cette époque, de départager ces hypothèses. En outre, on n'a pas encore établi fermement la continuité des chromosomes : comme on vient de le voir, Goldschmidt est partisan de la thèse selon laquelle ils se dissolvent et se réagrègent. Toutefois, malgré cette divergence fondamentale sur le support de l'hérédité, Goldschmidt admet deux des principes fondamentaux de la cartographie : le crossing-over et la linéarité. En revanche, la cartographie et la notion de « distance » qui l'accompagne ont pour lui une valeur métaphorique : les relations spatiales des cartes ne représentent pas des relations spatiales, mais des différences d'intensité entre les forces biochimiques associées aux gènes.

5.2 La critique de Bateson (1916)

La critique de William Bateson, mendélien de la première heure, est d'une tout autre nature. Contrairement à Goldschmidt, c'est un généticien authentiquement mendélien ; comme on l'a vu, la découverte des phénomènes de liaison partielle est autant son fait que celui de Morgan. On peut donc considérer sa critique comme une critique interne au mendélisme, ce qui est moins évident dans le cas de Goldschmidt.

Dans sa recension du *Mechanism of Mendelian Heredity*, Bateson s'oppose farouchement à la théorie chromosomique (contrairement à Goldschmidt). En revanche, il admet la pertinence de la construction de cartes génétiques, et accepte l'hypothèse de linéarité – pour laquelle, encore une fois, on ne dispose pas de preuve empirique. Son refus de la théorie chromosomique s'accompagne du rejet de la notion de crossing-over (c'est-à-dire de l'explication mécanique d'un échange de facteurs – quel que soit son support physique). Il la considère en effet comme faisant partie d'une grande machinerie d'hypothèses *ad hoc* qui sont destinées à faire tenir l'hypothèse chromosomique, plutôt que d'en être d'authentiques confirmations. Aucune preuve matérielle que les chromosomes se cassent effectivement pendant la méiose ne venant soutenir la théorie

de Morgan et de ses élèves, tout l'édifice de la théorie chromosomique lui paraît tenir sur du sable. Bateson se représente les gènes comme des charges, plutôt que comme des particules matérielles. Selon lui, cette représentation aurait l'avantage de fournir une explication du développement embryologique.

La position de Bateson permet de mettre en évidence, en regard de celle de Goldschmidt, que l'on peut accepter les principes mendéliens fondamentaux (ceux qui sous-tendent l'utilisation du symbolisme introduit par Mendel) sans pour autant adopter non seulement l'hypothèse chromosomique, mais encore une représentation des gènes comme des particules matérielles. Elle suggère en outre que l'hypothèse du crossing-over, qui semble très plausible à un regard rétrospectif, à la lumière des preuves cytologiques ultérieures, avait encore à cette époque un statut hautement hypothétique, et consistait en l'interprétation de données numériques sur une base observationnelle très limitée. En revanche, la pertinence de la construction de cartes génétiques – sans signification cytologique –, est admise avec l'hypothèse de linéarité, dont la formulation par Morgan était pourtant due à l'explication mécaniste du crossing-over. Notons (voir le tableau récapitulatif 2, page 525) que Goldschmidt, anti-mendélien, admet trois hypothèses du groupe de Morgan sur quatre, quand Bateson en rejette trois sur quatre.

5.3 Le débat avec Castle (1919) sur la linéarité

La critique de Castle (1919) est d'une nature encore différente des deux précédentes. Tout en acceptant plusieurs des hypothèses du groupe de Morgan, en particulier la théorie chromosomique, ainsi que le principe même de la cartographie génétique consistant à représenter les fréquences de recombinaison par des distances, Castle procède à une attaque en règle de la manière dont cette technique est mise en œuvre par le groupe de Morgan ; sa critique s'accompagne de la proposition d'un modèle différent de carte génétique. Le débat qu'elle engendre concerne donc la technique de cartographie elle-même. Dans ce qui suit, je présente les arguments avancés de part et d'autre ; j'en propose une analyse dans la section 6.

5.3.1 La critique de Castle (1919)

Contrairement à Bateson, Castle admet la théorie chromosomique³⁵, c'est-à-dire l'hypothèse selon laquelle les gènes sont localisés sur les chromosomes, sans que les modalités de cette « localisation » fassent l'objet d'une hypothèse claire. Il admet

³⁵ « Morgan a suggéré que ce qui attache ou relie deux caractères est le fait que leurs gènes se trouvent sur le même corps au sein du noyau de la cellule. Il suppose que ces corps sont les chromosomes. Il existe des preuves très fortes en faveur de cette conclusion. (Castle, 1919, p. 25)

également l'idée d'un échange entre portions de chromosomes (crossing-over cyto-logique ou chiasmatypie) et celle selon laquelle les emplacements des gènes au sein d'un groupe de liaison « ont une relation très définie et constante entre eux » (Castle, 1919, p. 25), justifiant par là-même l'ambition de développer une technique de carto-graphie permettant de prédire avec précision les phénomènes de recombinaison entre facteurs.³⁶

Sur cet accord de principe, Castle s'attaque à une des hypothèses fondamentales de la technique de cartographie génétique mise au point par Sturtevant, celle de l'*ordonnement linéaire* des gènes sur les chromosomes. En ce sens, la critique de Castle est diamétralement opposée à celle de Bateson, qui refuse l'hypothèse chromo-somique mais admet la linéarité au principe de la théorie du crossing-over, comprise comme théorie de l'échange de portions de groupes de liaison, non nécessairement identifiés aux chromosomes.

Les arguments de Castle contre l'hypothèse de linéarité Contre cette hy-pothèse, Castle avance des arguments de différentes sortes. Tantôt, comme le titre de l'article semble le suggérer (« L'ordonnement des gènes sur les chromosomes est-il linéaire ? »), Castle remet en question l'hypothèse mécanique selon laquelle les gènes sont effectivement attachés les uns aux autres le long des chromosomes, ces derniers ayant une forme de fils. Autrement dit, Castle soulève la question de savoir si le mécanisme du crossing-over *au niveau chromosomique* est réellement tel que le représentent les cartes de l'école de Morgan et le modèle du collier de perles. Tan-tôt, Castle affirme que, indépendamment de cette première question, qui concerne le niveau cytologique, les cartes du groupe de Morgan ne fournissent pas des prédic-tions satisfaisantes *au niveau génétique*, c'est-à-dire relatives aux données numériques obtenues par des expériences de croisement. Comme on va le voir, seuls ces derniers arguments sont présentés comme décisifs par Castle ; en proposant son propre modèle de carte, sa préoccupation majeure est en effet d'améliorer l'exactitude et le pouvoir prédictif de cet outil, sans qu'il prétende rien affirmer sur le mécanisme chromoso-mique sous-jacent.

La critique de Castle s'ouvre et se ferme par des considérations concernant la structure des chromosomes. Avant de procéder à un examen détaillé des défauts du modèle de carte proposé par le groupe de Morgan, Castle affirme en effet que l'hypothèse fondamentale de la linéarité est très peu plausible en tant qu'hypothèse portant sur la morphologie réelle des chromosomes : « Il est douteux [...] qu'une

³⁶ « Les suppositions ont largement prouvé leur utilité comme hypothèses de travail, puisque qu'il s'est avéré possible, connaissant certaines valeurs de crossing-over, d'en prédire d'autres à un bon degré de précision. » (Castle, 1919, p. 26)

molécule organique complexe puisse avoir la forme simple d'un fil. » (Castle, 1919, p. 26) Une fois cet argument avancé, Castle passe aux arguments relatifs à la valeur prédictive du modèle au niveau génétique ; ces arguments occupent l'essentiel de son propos. Ce n'est qu'à la toute fin de l'article que, ayant présenté son propre modèle de carte, construit pour répondre à des exigences d'ordre prédictif au niveau génétique, il suggère des hypothèses relatives au mécanisme cytologique qui pourrait expliquer les phénomènes de crossing-over.

Les arguments centraux de Castle concernent la recherche de la meilleure *présentation des données*. Il est important de souligner que les données qu'il utilise sont celles de Morgan et Bridges (1916). À aucun moment, ces données ne sont contestées par Castle : l'essentiel de son propos consiste à comparer deux manières de les présenter, celle du groupe de Morgan et la sienne. Par cette comparaison, Castle prétend mettre en évidence la supériorité de son modèle sous deux aspects principaux : il permet des prédictions plus exactes, et il ne repose sur aucune hypothèse *ad hoc*.

La technique de cartographie du groupe de Morgan souffre en effet, selon Castle, de deux défauts majeurs, tous deux liés à l'hypothèse d'un ordonnancement linéaire des gènes. Tout d'abord, cette hypothèse repose sur un arsenal coûteux d'hypothèses *ad hoc* destinées à s'accommoder des exceptions à l'additivité et par conséquent à la linéarité³⁷ de la fonction de distance par rapport au nombre de recombinaisons observées. « Pour rendre compte de cette divergence, Morgan a adopté des hypothèses subsidiaires, d'interférence, de "double crossing-over", etc. » (Castle, 1919, p. 26)

Or, selon Castle, rien ne justifie de tels frais ; bien au contraire, maintenir l'hypothèse de linéarité conduit à des prédictions incohérentes. En effet, la technique de Sturtevant qui consiste à construire la carte d'un chromosome en se basant sur la somme des plus petites distances (pour éviter les difficultés liées aux manquements à l'additivité pour les grandes distances, dues selon Sturtevant à des crossing-over doubles) conduit, selon Castle, à des résultats impossibles : la carte du chromosome *X* proposée par Morgan et Bridges (1916, voir figure 71) présente des distances de plus de 50 unités cartographiques. Or, une unité de distance correspond à 1% de crossing over ; 50% est donc le maximum possible, atteint dans les cas d'assortiment indépendant.

³⁷ Une mise en garde s'impose concernant les deux usages distincts du terme « linéarité » dans la présentation de la critique de Castle. Dans un cas, la linéarité désigne l'hypothèse de l'ordonnancement physique des gènes le long des chromosomes, supposés avoir la forme d'un fil, comme des perles sur un collier. Dans l'autre, elle désigne la linéarité de la fonction de distance par rapport au nombre de crossing-over effectifs. Castle refuse l'hypothèse de la linéarité de l'ordonnancement des gènes, ce qui lui permet de proposer une fonction de distance qui soit linéaire, contrairement à ce qui est le cas chez Morgan.

La carte de Morgan pour le chromosome sexuel place cinq gènes sur vingt-neuf à des distances de 55 et de 66 de l'extrémité du chromosome, où le facteur jaune est situé [facteur 0]. Un instant de réflexion suffit à montrer que ces relations sont impossibles, puisqu'un pourcentage de crossing-over supérieur à 50 est absurde. Si *A* et *B* s'assortissent entièrement indépendamment, sans aucune liaison, comme ils le feraient dans un cas ordinaire d'hérédité mendélienne où il n'y a pas de liaison, le nombre de crossing-over et de non crossing-over serait le même, 50% chacun. Toute valeur de crossing-over inférieure à 50% manifeste une liaison. *Une valeur de crossing-over supérieure à 50% ne peut pas exister*.³⁸ En effet, soit il y a liaison, soit il n'y a pas liaison. Or l'absence de liaison revient à 50% de crossing-over, et la liaison revient à moins de 50% de crossing-over. Donc une valeur plus grande que 50% ne peut pas se produire. (Castle, 1919, p. 28)

Cette analyse repose manifestement sur un contresens de Castle concernant la notion de crossing-over, due à l'ambiguïté du terme mentionnée précédemment. En effet, l'emploi du terme « crossing-over » par Castle correspond à celui de crossing-over *observé*, c'est-à-dire de recombinaison effective. Or, les grandes distances cartographiques ne correspondent pas, en raison, précisément, de l'hypothèse de doubles crossing-over, à des probabilités de recombinaison effective. Castle, refusant, comme on l'a vu, ces hypothèses qu'il juge *ad hoc*, ne fait aucune différence entre crossing-over *réel* et recombinaison effective. Il va même, assez étrangement, jusqu'à utiliser le terme « crossing-over » pour désigner tous les cas où deux gènes ne sont pas hérités ensemble, y compris quand ces gènes n'appartiennent pas à un même groupe de liaison ; c'est ce qui lui permet de dire que dans les cas d'assortiment indépendant, il y a le même nombre de « crossing-over » que de « non crossing-over » (c'est-à-dire d'hérédité commune des deux gènes). Je reviendrai sur ce contresens de Castle lors de mon analyse du débat, à la section 6.

L'hypothèse de la linéarité – en plus d'être peu plausible en tant qu'hypothèse sur la morphologie des chromosomes – s'avère donc, selon Castle, non seulement coûteuse, mais encore nocive ; rien n'en justifie le maintien. Y renoncer fait gagner en exactitude et en simplicité ; c'est sur l'abandon de cette hypothèse qu'est fondée la méthode de cartographie proposée par Castle.

Le modèle de Castle Puisque les fréquences de recombinaison entre les facteurs les plus éloignés sont inférieures à la somme des fréquences de recombinaison des facteurs plus proches, la représentation graphique des distances (définies comme des probabilités de recombinaisons) ne peut pas être unidimensionnelle. « Au lieu d'un ordonnancement linéaire, on peut supposer que l'ordonnancement des gènes n'est pas

³⁸Souligné par Castle.

linéaire, et que les gènes les plus proches ne sont pas sur la même ligne que les plus éloignés. » (Castle, 1919, p. 28)

Castle propose alors son propre modèle de présentation des données obtenues par Morgan et Bridges (1916), que ces derniers présentaient sur une carte linéaire (figure 71) ; le modèle de Castle répond à une exigence essentielle : la distance, quelle qu'en soit la définition, doit être une fonction linéaire de la fréquence de recombinaison. Il s'agit bien d'une carte génétique au sens de Sturtevant, puisqu'elle opère une conversion des fréquences de recombinaison observées en une mesure de distance ; cependant, abandonnant l'hypothèse de linéarité, la technique de construction n'est pas la même : toutes les fréquences de recombinaison, même les plus importantes (c'est-à-dire concernant les facteurs les plus éloignés), sont à prendre en compte pour mesurer les distances. Prenons par exemple trois facteurs, A , B et C . Si l'on veut les présenter sur un diagramme de telle manière que les distances qui les séparent soient proportionnelles à leurs fréquences de recombinaison deux à deux, le fait que la fréquence de recombinaison de A et C soit inférieure à la somme des fréquences de recombinaison de A et B et de B et C implique que A , B et C soient les sommets d'un triangle. Le modèle de Castle prend donc la forme d'une figure tridimensionnelle (figure 72 page 520) :

Si l'ordonnement des gènes n'est pas linéaire, comment le caractériser ? Cette question m'a conduit à proposer une présentation graphique de la relation indiquée par les données de Morgan et Bridges³⁹, mais, constatant que cette méthode n'était pas satisfaisante, j'ai effectué une reconstruction en trois dimensions, qui s'est avérée très satisfaisante. (Castle, 1919, p. 28)

Castle présente son modèle comme découlant naturellement de la présentation des données de Morgan et Bridges (1916) de la manière la plus simple :

La seule hypothèse qui sous-tend cette reconstruction est la plus fondamentale de Morgan, celle selon laquelle les distances entre les gènes sont proportionnelles aux pourcentages de crossing-over. Je rejette la seconde hypothèse, celle selon laquelle les gènes distants sont en réalité plus éloignés les uns des autres que ce qui est indiqué par les données expérimentales, pour la raison expliquée précédemment, que cela implique des relations impossibles. Prenant donc les données exactement telles qu'elles sont, il est possible d'opérer une reconstruction complète et entièrement cohérente de l'architecture des chromosomes sexuels à partir des relations de liaison entre leurs gènes. (Castle, 1919, p. 28)

Castle présente donc son modèle comme une mise en œuvre plus pure de l'hypothèse initiale de Morgan selon laquelle la fréquence de recombinaison est un indice de la distance : sans y ajouter aucune hypothèse supplémentaire, il parvient à construire

³⁹(Morgan et Bridges, 1916)

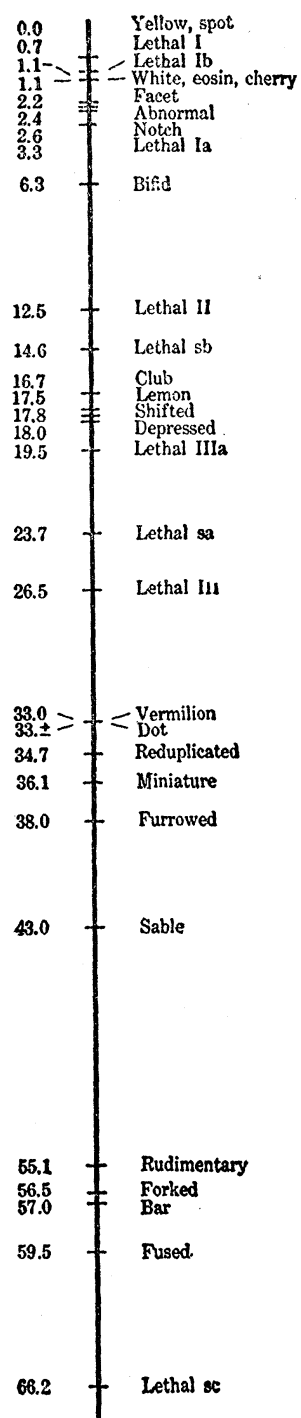


FIG. 71 — Carte du chromosome X de Morgan et Bridges (1916, p. 23), reproduite par Castle (1919, p. 23)

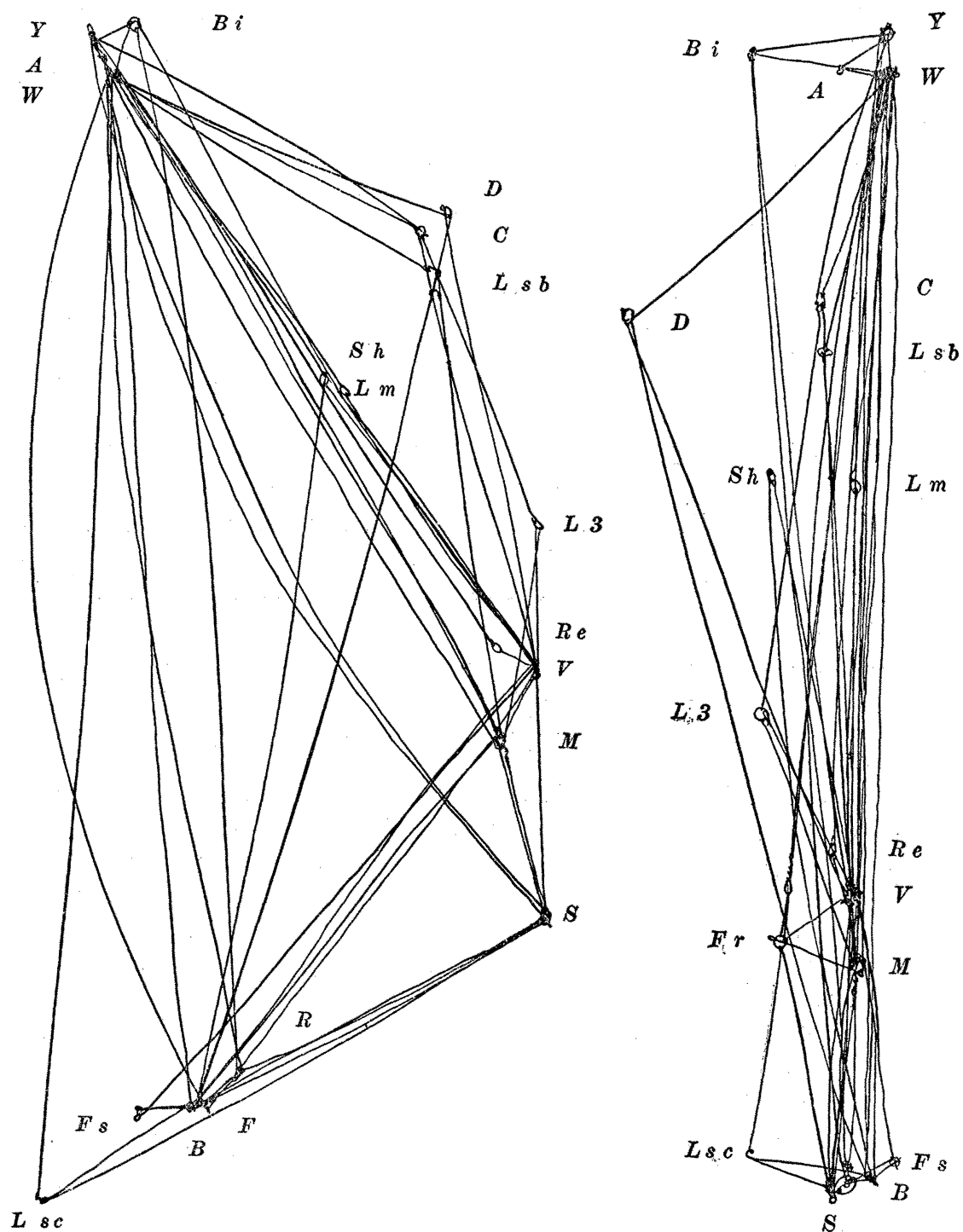


FIG. 72 – Le modèle du « piège à rat » de Castle (1919, p. 29) sous deux perspectives différentes (la figure de droite est obtenue par une rotation de 90°).

un modèle plus cohérent, mieux en accord avec les données obtenues par Morgan et Bridges (1916) eux-mêmes, et qui plus est beaucoup plus simple que celui de ses adversaires. En effet, il est la mise en œuvre d'une seule et unique hypothèse, celle de la proportionnalité de la distance à la fréquence de recombinaison. L'abandon de l'hypothèse de la linéarité s'accompagne de celui de toutes les hypothèses subsidiaires mentionnées précédemment, comme celles du double crossing-over et de l'effet d'interférence.⁴⁰ C'est la raison pour laquelle, comme je l'ai noté précédemment, les expressions « recombinaison observée » et « crossing-over » deviennent synonymes : tout crossing-over est nécessairement manifeste (voir mon analyse de cette ambiguïté page 497).

Récapitulons rapidement l'argumentation de Castle. Premièrement, il est selon lui fort peu probable que les chromosomes soient de simples fils : cet argument, relatif au niveau cytologique, a simplement pour rôle, dans l'économie de l'article de Castle (1919), de montrer que les raisons initiales de supposer la linéarité sont faibles. Deuxièmement, l'argument central, et présenté comme décisif par Castle, est que l'hypothèse du groupe de Morgan conduit à construire un modèle incohérent qui produit parfois de mauvaises prédictions au niveau génétique. Enfin, le modèle que propose Castle présente le grand avantage d'être plus *simple* que celui de Morgan : il ne nécessite aucune hypothèse *ad hoc*.

Avant d'en venir aux réponses du groupe de Morgan, arrêtons-nous un instant sur les arguments concernant la structure des chromosomes. Castle admet, on l'a vu, l'hypothèse chromosomique. Cependant, celle-ci ne signifie rien d'autre que le fait que les gènes sont situés sur les chromosomes, et que les phénomènes génétiques (ségrégation, assortiment indépendant, liaison partielle) doivent pouvoir recevoir une explication au niveau cytologique. L'imprécision des observations cytologiques laisse le plus grand flou sur les mécanismes à l'œuvre au sein des chromosomes et sur leur morphologie. En conséquence, la recherche d'une explication cytologique ne peut être que spéculative, et les arguments qui concernent la structure des chromosomes sont explicitement présentés comme annexes. Autrement dit, la carte de Castle ne prétend pas être autre chose qu'une machinerie analogue à celle d'un planétarium permettant de reproduire le mouvement apparent des astres. Comme le suggère Wimsatt (1987), Castle présente sa carte comme un modèle « phénoménologique » au sens de Cartwright (1983) : il rend compte des données mais ne prétend pas les expliquer. Cependant, après avoir défendu son modèle de carte en raison de sa plus grande

⁴⁰ Le maintien de l'hypothèse de linéarité et de celle des crossing-over doubles repose, selon Castle, sur une forme de circularité : « Les données expérimentales prouvent que le double crossing-over doit se produire si l'ordonnement des gènes est linéaire. » (Castle, 1919, p. 30) Or le double crossing-over est une hypothèse *ad hoc* convoquée pour maintenir la linéarité.

efficacité prédictive, Castle ajoute que cette valeur prédictive indique qu'il est assez probable que ce modèle représente des relations réelles au sein du chromosome : « Ces prédictions doivent être plus proches de la vérité que celles fondées sur une carte linéaire si, comme je l'ai suggéré, l'arrangement *n'est réellement pas linéaire*⁴¹. » Ce que semble vouloir dire Castle est que la valeur prédictive de la représentation tridimensionnelle des probabilités de recombinaison suggère que les mécanismes à l'œuvre au niveau cytologique doivent correspondre, d'une manière ou d'une autre, aux relations représentées dans le diagramme. Autrement dit, plus un modèle est prédictif, plus il est probable, selon Castle, que cette « machinerie » – en l'occurrence, les règles de construction et d'interprétation de la carte selon le modèle qu'il en propose – corresponde aux mécanismes réels.

Par « mécanisme », cependant, Castle n'entend pas suggérer que les relations représentées par la carte soient effectivement des relations spatiales :

On me demandera peut-être ce que cette reconstruction signifie. Présente-t-elle la forme réelle du chromosome, ou du moins cette partie du chromosome où les variations génétiques ont lieu ? Où est-ce simplement la représentation symbolique de forces moléculaires ? Ces questions ne peuvent pas, en l'état actuel de nos connaissances, recevoir de réponse. [...] Un modèle qui répond de façon véridique à des questions doit être la présentation véridique de relations réelles, quand bien même on ne sait pas si ce sont des relations spatiales ou dynamiques.

(Castle, 1919, p. 30)

Il convient donc de distinguer entre deux choses différentes. D'une part, Castle cherche à défendre son modèle sur la seule base de sa valeur prédictive, indépendamment de la question de la relation entre la carte et les mécanismes à l'œuvre au niveau cytologique. D'autre part, une fois reconnue la grande probabilité que la carte soit une représentation « véridique de relations réelles », en raison même de sa grande valeur prédictive, Castle prône un agnosticisme quant à la nature de ces relations.

Dans les termes définis au chapitre 7, la question de Castle est la suivante : la carte est-elle effectivement un schéma (tridimensionnel), dont les relations spatiales représentent des relations spatiales, ou bien est-elle un diagramme, au sein desquelles les relations spatiales représentent des relations non spatiales, comme par exemple des relations entre forces ? La « distance » génétique correspond-elle, de manière approximative, à une distance physique, ou bien la notion de distance est-elle métaphorique ? La fin de l'article de Castle (pp. 31-32) penche en faveur de la dernière solution ; en effet, Castle suggère une explication chimique du phénomène de crossing-over, qu'il décrit « comme le remplacement d'une base chimique par une autre au sein d'une molécule organique complexe » et dont il affirme qu'« il semble hautement probable

⁴¹ Je souligne.

que ce soit sa vraie nature. » (p. 32)

5.3.2 Les réponses de Sturtevant *et al.* (1919) et Muller (1920)

Le groupe de Morgan réagit à la critique de Castle en plusieurs articles (en particulier Sturtevant *et al.*, 1919; Muller, 1920), qui, pris ensemble, répondent à chacun des arguments de Castle. En résumé, les réponses du groupe de Morgan montrent que, à supposer même que Castle utilise correctement les données (qui sont aussi les leurs) – ce qui n'est pas le cas –, ces données ne soutiendraient pas son modèle. En outre, l'inclusion de nouvelles données ferait apparaître de manière encore plus manifeste l'inadéquation du modèle. Enfin, aucun des mécanismes supposés par Castle dans son article pour rendre compte de son modèle n'aurait selon eux les effets désirés. Je ne présente ici que certains des arguments du groupe de Morgan, avant de proposer, à la section 6, une analyse de ce débat.

La réponse de Sturtevant *et al.* (1919) accorde une importance assez secondaire aux arguments de Castle concernant la structure des chromosomes et affirment simplement que, contrairement à ce que suggère Castle, les chromosomes ne sont pas des molécules organiques ; ils refusent l'analogie entre crossing-over et réaction de substitution organique. En ce qui concerne l'exactitude des prédictions, Sturtevant *et al.* (1919), ainsi que Muller (1920) donnent de nombreux arguments pour montrer que le modèle de Castle, contrairement à ce qu'il prétend, ne rend pas correctement compte des données. Le principal argument de Muller (1920) dans ce sens consiste à montrer que la distance entre les facteurs les plus éloignés (ceux qui sont séparés par 4 facteurs), telle qu'elle est prédite par le modèle de Castle, est trop grande (elle est de 49,3 alors que la fréquence de recombinaison observée est de 41,6). Or, si l'on « plie » la figure afin de rendre les grandes distances correctes, les courtes distances deviennent trop courtes, et le modèle devient encore plus inexact.⁴²

En outre, Sturtevant *et al.* (1919) soulignent le fait que la remarque de Castle à propos des valeurs de crossing-over supérieures à 50% repose sur une mésinterprétation de la notion de distance. Reproduisant l'argument de Castle cité ci-dessus⁴³, ils affirment que « la conclusion de ce curieux syllogisme dépend entièrement de la définition de la liaison contenue dans la deuxième partie de la seconde prémisse. » (Sturtevant *et al.*, 1919, p. 170) La phrase en question est : « la liaison revient à moins de 50% de crossing-over. » J'ai déjà souligné l'usage particulier (et hétérodoxe

⁴²Pour une analyse détaillée des arguments de Muller (1920), voir (Wimsatt, 1987, 1992).

⁴³« [...] soit il y a liaison, soit il n'y a pas liaison. Or l'absence de liaison revient à 50% de crossing-over, et la liaison revient à moins de 50% de crossing-over. Donc une valeur plus grande que 50% ne peut pas se produire. »

du point de vue de l'école de Morgan) que Castle fait de la notion de crossing-over et de sa relation à celle de liaison. En opérant une distinction entre les cas de liaison (partielle ou non) et les cas d'absence de liaison qu'il fait équivaloir à la distinction entre 50% de crossing-over et moins de 50%, il fait une pétition de principe : ainsi définie, la notion de crossing-over rend impossible l'idée même de crossing-over double. En effet, un crossing-over double, pour Castle, est une absence de crossing-over, donc une liaison. La notion de crossing-over, pour le groupe de Morgan, ne peut pas être appliquée aux gènes qui ne sont jamais liés. 50% de crossing-over, tels que Castle les définit, ne comptent pas comme des crossing-over pour le groupe de Morgan, mais pour une absence de liaison. En revanche, une fois les groupes de liaison établis, et l'hypothèse de linéarité admise, les grandes distances sur un même groupe de liaison peuvent tout à fait conduire à admettre l'existence de crossing-over multiples, qui sont comptabilisés comme des distances supérieures à 50%. Les distances ne correspondent pas aux crossing-over observés qui, effectivement, soit ont lieu soit n'ont pas lieu, mais aux crossing-over réels.

Pour finir, je voudrais m'arrêter un instant à un des aspects de la réponse de Muller (1920). Après avoir souligné le contresens de Castle sur la notion de crossing-over⁴⁴, il affirme que la linéarité des gènes est une hypothèse indispensable pour expliquer les phénomènes génétiques, *indépendamment de la question de la réalité physique correspondant à cette linéarité* :

Que l'on considère ou non les facteurs comme localisés sur un fil matériel réel, on doit, sur la base de ces découvertes, admettre que les forces qui les lient ensemble – qu'elles soient physiques, « dynamiques » ou transcendantes – sont d'une nature telle que chaque facteur est lié, au cours de la ségrégation, uniquement à deux autres facteurs – de telle sorte que le groupe entier, considéré dynamiquement, est une chaîne. [...] rien n'est impliqué quant à l'ordonnement physique des gènes quand on emploie les termes « séries linéaires », « distance », etc. (Muller, 1920, p. 101)

Avec une certaine ironie, Muller suggère donc que, à supposer même que les forces qui lient les gènes entre eux soient de la nature supposée par Castle, la linéarité doit être maintenue, comme l'hypothèse selon laquelle les gènes sont liés *un à un*. Muller semble donc suggérer que le modèle du collier de perles peut être interprété de façon analogique – le collier n'est pas forcément à un « fil matériel réel » – mais que sa valeur explicative réside dans l'hypothèse de linéarité, physique ou non. Sans

⁴⁴ « [Nous n'avons] jamais prétendu, avec la théorie de la liaison linéaire, que les pourcentages de crossing-over étaient effectivement proportionnels à des distances cartographiques : ce qui a été affirmé est que ces pourcentages de crossing-over sont calculables à partir de la distance cartographique [...]. » (Castle, 1919, p. 98)

cette hypothèse, tous les phénomènes génétiques observés jusqu'ici deviennent inexplicables. L'argument de Muller concerne donc à la fois la valeur prédictive du modèle l'explication mécanique qu'il apporte : parce qu'il est impossible de rendre compte des données sans les présenter de manière linéaire, on est obligé de supposer que, réellement, les gènes sont ordonnés linéairement. En revanche, Muller prétend, ici, rester agnostique quant à la nature physique de cette « chaîne » formée par les gènes.

Il ajoute cependant, quelques lignes plus loin, que « quand les diverses conditions qui doivent être remplies au moment de la ségrégation sont prises en considération, toute autre explication de ces découvertes de liaison linéaire qu'un arrangement des gènes sur une ligne physique, spatiale, apparaît hasardeuse et fantaisiste. » (Muller, 1920, p. 101) Autrement dit, une fois admise l'hypothèse de linéarité, on n'est plus vraiment justifié à résister à son interprétation spatiale. Muller présente ici son engagement en faveur de l'hypothèse chromosomique en tant qu'explication mécanique des phénomènes de crossing-over comme une conséquence de son souci de rendre compte au mieux des données empiriques.

Bien que l'histoire ait donné raison au groupe de Morgan, on peut cependant considérer, avec Castle, que la défense par Muller du crossing-over et de la linéarité reposent à l'inverse sur un acte de foi en faveur d'une explication mécaniste, dont les diverses preuves ne sont réellement convaincantes que pour les « convertis ». On va voir, dans l'analyse qui suit, que le jeu entre engagement en faveur d'une explication mécanique et souci de fournir un modèle prédictif est plus compliqué que ne le laissent entendre les propos explicites des protagonistes du débat, qu'il s'agisse de Castle ou de Muller.

	Chrom	COcyto	COgéné	Lin
Goldschmidt	X	0	X	X
Bateson	0	0	0	X
Castle	X	X	X	0

TAB. 2 – Les positions respectives de Goldschmidt, Bateson et Castle sur les hypothèses chromosomique (Chrom), du crossing-over cytologique (COcyto), du crossing-over génétique (COgéné) et de la linéarité (Lin). « X » signifie acceptation ; « 0 » refus.

6 Interprétation de ces débats : que représentent les cartes génétiques ?

La présentation du débat entre Castle (1919) et le groupe de Morgan a fait apparaître la chose suivante : de part et d'autres sont avancés des arguments relatifs à la plausibilité des explications mécaniques apportées par les différents modèles de cartes proposés et des arguments relatifs à l'efficacité prédictive de ces modèles. De ce point de vue, l'opposition suggérée par Wimsatt (1987) entre un Castle guidé par un impératif « opérationnaliste » et un Morgan « réaliste » est partiellement erronée⁴⁵ : les convictions de Castle en faveur d'une représentation chimique des phénomènes de crossing-over jouent un rôle dans le choix de son modèle⁴⁶ ; inversement, Morgan a plusieurs fois clamé son agnosticisme en ce qui concerne la nature matérielle des gènes (voir chapitre 8, page 439).

Notons cependant que, aussi bien chez Sturtevant *et al.* (1919) et Muller (1920) que chez Castle (1919), les arguments véritablement présentés comme décisifs sont les arguments relatifs à l'exactitude des prédictions génétiques du modèle. Quand une référence est faite à la plausibilité du modèle en tant que représentation des mécanismes chromosomiques, c'est-à-dire en tant qu'explication cytologique et non plus seulement génétique, elle est explicitement présentée comme une remarque annexe : en ce sens, chacun des protagonistes de ces débats en reste au niveau strictement génétique, et aucun des élèves de Morgan n'affirme que le modèle du collier de perles est autre chose qu'une analogie, ni, par conséquent, que les cartes génétiques représentent effectivement des relations spatiales au sein du chromosome.⁴⁷

Je voudrais à présent montrer que les engagements de chacun de ces théoriciens en faveur d'une explication mécanique, chimique, ou autre, des phénomènes génétiques décrits par les cartes, jouent un rôle bien plus important – chez les élèves de Morgan

⁴⁵Wimsatt (1987) en conclut que ce débat est un cas particulièrement probant de la « supériorité d'un programme de recherche mécanique ou réaliste sur un programme opérationnaliste ou instrumentaliste ».

⁴⁶Lindley Darden (1991) suggère ainsi que la position de Castle est probablement due à sa formation de chimiste, et à une préférence prononcée pour les modèles de chimie organique, comme en témoigne sa proposition, quelques années avant, d'un premier modèle bidimensionnel.

⁴⁷Comme me l'a fait remarquer Patrice David, c'est même la volonté d'être des généticiens « les plus purs possibles » qui permet aux morganiens de rendre le parallélisme avec le comportement des chromosomes patent : le rapprochement entre génétique et cytologie apparaît ainsi comme celui de deux lignes de recherche indépendantes, ce qui lui donne d'autant plus de force de conviction. C'est d'ailleurs ce que suggèrent les propos de Morgan (1928, p. 44), que j'ai rapportés au chapitre 8 (page 448), sur la coïncidence entre les recherches génétiques et les découvertes cytologiques qui n'ont jamais été accomplies, selon lui, avec « un biais génétique ».

comme chez Castle – que leurs remarques explicites à ce propos ne le laissent entendre. Les arguments relatifs à la structure des chromosomes ne viennent pas simplement s'ajouter aux arguments qui consistent à montrer que l'adoption d'un certain modèle de carte conduit à de meilleures prédictions : *les engagements qu'ils révèlent fondent la compréhension même que les différents théoriciens ont de ces cartes et l'utilisation qu'ils en font*. C'est donc par une analyse, en plus de leurs propos explicites, de la manière dont chacun des protagonistes du débat construit, interprète et utilise les cartes génétiques, que je prétends rendre compte du jeu conjugué de ces préoccupations relatives à la prédiction des phénomènes génétiques et des engagements théoriques de chacun en faveur d'une explication au niveau cytologique.⁴⁸

Pour cela, je propose, à la double lumière de la présentation des débats entre le groupe de Morgan et de ses adversaires et de l'analyse des types de représentation introduite au chapitre 7, d'examiner la question suivante : que représentent les cartes génétiques ? On a vu, au chapitre 7 (section 1.3), que différents types de représentation (pictural, schématique, diagrammatique) semblent adaptés à l'expression de différents types de contenus (représentation des relations spatiales au sein d'une entité concrète, représentations d'une relation entre variables, etc.) Aussi, je me propose d'aborder la question de savoir ce que représentent les cartes génétiques sous l'angle de la question suivante : à quel type de représentation appartiennent les cartes génétiques ?

6.1 Les cartes génétiques : schémas ou diagrammes ?

Rappelons brièvement les définitions que j'ai proposées au chapitre 7 (sections 1.3.3 et 1.3.4) des notions de schémas et de diagrammes. Un *schéma* – ou représentation schématique – est une représentation au sein de laquelle les relations spatiales tiennent lieu de relations spatiales ; plus précisément, ces représentations respectent les relations topologiques de leur cible, qui est une entité matérielle concrète. Une telle représentation peut être obtenue par la schématisation d'une représentation picturale comme, par exemple, le micrographe d'un chromosome, auquel cas elle représente la morphologie de l'objet représenté (le chromosome) ; elle peut également être un simple dessin, en principe superposable (moyennant certaines déformations, comme

⁴⁸Ces engagements sont très probablement le fruit de l'influence conjuguée de la formation des chercheurs, de leurs connaissances d'arrière-plan et de leurs talents personnels, comme je l'ai suggéré à plusieurs reprises (Lindley Darden, 1991, accorde une grande importance à ces facteurs). Cependant, à moins que ces préférences ne soient explicitées, ces influences sont très difficiles à mesurer pour l'historien. C'est la raison pour laquelle, plutôt qu'une reconstruction hypothétique des croyances et des préférences de chacun, je me concentre, dans mon analyse, outre les propos explicites des protagonistes, sur la manière dont chacun construit, interprète et utilise les cartes génétiques.

lorsque l'on représente la France par un hexagone ou un chromosome par un bâtonnet) à une représentation picturale. Le modèle du collier de perles se présente ainsi sous la forme d'un schéma de la morphologie hypothétique – et non observée – des chromosomes.

Un *diagramme* est également une représentation spatiale, au sens où les relations spatiales en son sein dénotent des propriétés de son référent⁴⁹ ; cependant, ces relations ne représentent pas des relations spatiales. Elles peuvent représenter des relations temporelles, des relations causales, des relations hiérarchiques, etc. On peut qualifier ces représentations de « métaphoriques », au sens où elles représentent des relations non spatiales par des relations spatiales, comme lorsque l'on parle de la « distance » temporelle entre deux événements. Les graphes, qui représentent la relation entre différentes variables, appartiennent à la catégorie des diagrammes.

Après avoir présenté les arguments qui invitent à considérer les cartes génétiques comme des schémas dont les relations spatiales représentent des relations spatiales au sein des chromosomes, je montrerai que ces arguments ne semblent pas résister à l'analyse, et que les cartes génétiques doivent bien plutôt être considérées comme des diagrammes, et plus précisément des graphes, présentant spatialement des données statistiques proprement génétiques, qui ne correspondent pas nécessairement à des distances spatiales. Cependant, à la section 6.2, on verra que cette dernière conclusion n'est elle-même pas satisfaisante, et que la manière dont les théoriciens du groupe de Morgan construisent et utilisent les cartes génétiques implique dans une certaine mesure de les interpréter comme des schémas.

6.1.1 Les cartes génétiques comme représentations schématiques de chromosomes

Comme je l'ai suggéré précédemment (page 510), la confirmation de l'hypothèse chromosomique apportée par Painter (1934a), qui permet de projeter la carte génétique obtenue par des méthodes mendéliennes sur la carte cytologique obtenue par la technique de la chambre claire (figures 70 et 73), peut être considérée comme une démonstration de la nature *schématique* des cartes génétiques. En effet, les cartes cytologiques sont typiquement des représentations picturales. En conséquence, la mise en correspondance de la carte cytologique (picturale) et de la carte génétique semble conférer *ipso facto* à la carte génétique le statut de schéma. Quand bien même la métrique des deux cartes n'est pas la même (la distance génétique ne correspond pas à la distance physique), les rapports topologiques de la carte cytologique sont conservés par la carte génétique (l'ordre relatif des gènes, sinon leur éloignement relatif). Le fait

⁴⁹Voir chapitre 7, sections 1.3.1 et 1.3.4.

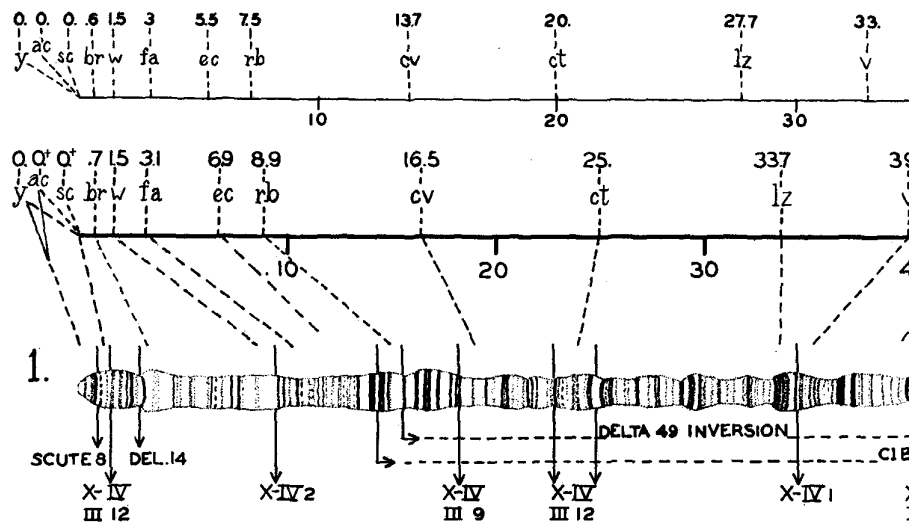


FIG. 73 – Moitié gauche des cartes génétique et cytologique du chromosome X de la drosophile. Tout en haut est présentée la carte génétique avant sa mise en correspondance avec la carte cytologique. Juste en dessous, la carte génétique est adaptée et projetée sur la carte cytologique (Painter, 1934a, p. 452).

que la mise en correspondance des deux cartes permette, en outre, de *corriger* (voir figure 73) certains éléments de la carte génétique est un symptôme supplémentaire de la nature schématique de cette dernière.

Comme on l'a vu, l'invention des cartes génétiques est due à l'adoption de l'hypothèse chromosomique. Ce que confirme la mise en correspondance des deux cartes, plus que l'hypothèse chromosomique proprement dite, c'est la dimension schématique des cartes génétiques, c'est-à-dire le fait que les relations spatiales au sein de la carte correspondent bien à des relations spatiales dans l'objet représenté. Autrement dit, la projection de la carte génétique sur la carte cytologique permet de montrer que le nom de « carte » n'est pas métaphorique : malgré certaines distorsions des distances, l'ordre relatif des gènes sur les chromosomes est bien restitué par la carte génétique. Les quatre cartes génétiques de *Drosophila* correspondent bien aux quatre chromosomes de cette espèce. En ce sens, les cartes tirées des données obtenues par expériences d'hybridation semblent bien être, au sens défini ci-dessus, des *schémas* de chromosomes. Les appeler « cartes chromosomiques », comme c'est souvent le cas, ne semble pas abusif : avant la confirmation cytologique, la construction de ces cartes a bien permis aux généticiens de faire des découvertes sur la morphologie des chromosomes.

6.1.2 Les cartes génétiques comme présentation graphique de données numériques

Cependant, plusieurs arguments peuvent être avancés en faveur de la position inverse, selon laquelle les cartes génétiques ne sont pas, à strictement parler, des *cartes*, mais plutôt des graphes (ou diagrammes) présentant spatialement des données statistiques ; ce ne sont des cartes qu'en un sens métaphorique. Dans cette perspective, la mise en correspondance des cartes cytologique et génétique, si elle confirme bien l'hypothèse chromosomique, ne suffit pas à affirmer que les cartes génétiques *représentent des chromosomes*. Autorisons-nous un instant à faire de l'histoire contrefactuelle : si l'hypothèse chromosomique n'avait pas été confirmée, l'utilité et la pertinence de la construction de cartes génétiques n'aurait certainement pas été annulée, comme en témoigne par exemple la position de William Bateson, présentée ci-dessus. En ce sens, les cartes génétiques appartiennent au domaine du mendélisme pur, et non à celui de la cytologie.

Tout d'abord, les cartes génétiques sont construites à partir d'une technique purement mendélienne. Cette technique consiste en des expériences d'hybridation qui permettent d'obtenir des données statistiques concernant la probabilité d'une recombinaison entre différents facteurs. Le tableau de la figure 56 et le graphe de la figure 57 (page 493) contiennent strictement les mêmes informations. En ce sens, les cartes peuvent être décrites comme des *extensions du symbolisme mendélien*, au même titre que les carrés de Punnett : les points de la carte représentent des unités hypothétiques (les gènes ou facteurs mendéliens) sans qu'aucune hypothèse sur la nature de ces facteurs soit explicitement présentée. Leurs relations spatiales sur la carte représentent des mesures de probabilité. Leur construction et leur interprétation sont indépendantes du fait qu'à la linéarité de la carte corresponde réellement la linéarité du chromosome, exactement de la même manière que les notions génétiques de groupe de liaison et de recombinaison de gènes sont indépendantes conceptuellement de celle de chiasmatypie (crossing-over cytologique). Comme je l'ai signalé, la distance génétique ne correspond pas à la distance physique sur le chromosome : les variations du taux de recombinaison sont telles que la probabilité de recombinaison n'est pas constante par micromètre de longueur du chromosome.⁵⁰ Les drosophiles mâles ont les mêmes chromosomes que les femelles (sauf le *X*) mais recombinent beaucoup moins ; en conséquence, ils n'ont pas la même carte génétique. Même si l'on parvenait à identifier un point précis visible sur les chromosomes pour chaque

⁵⁰Notons d'ailleurs que, comme me l'a signalé Patrice David, le cadre moléculaire ne fait pas non plus correspondre une distance physique précise à la distance génétique : la probabilité du taux de recombinaison n'est pas non plus constante par nucléotide d'ADN.

gène, l'objet principal des cartes génétiques (prédire les pourcentages de recombinaison) ne pourrait pas être atteint par des techniques cytologiques.

De ce point de vue, il semble que le contenu des cartes soit à strictement parler le même que celui des tableaux présentant les données obtenues par expérience d'hybridation. La présentation graphique de ces données permet de tirer des inférences de façon plus rapide et moins coûteuse que ne le permettent les données en tableau sur les probabilités de recombinaison des différents facteurs : ces probabilités sont rendues immédiatement visibles par les plus ou moins grandes distances sur la carte. Cependant, les inférences que l'on peut tirer de la carte concernent des prédictions numériques de même nature que celles que l'on peut tirer de la considération des résultats des expériences et des données à partir desquelles la carte a été construite. Autrement dit, elles permettent de visualiser ces données, et non les chromosomes.

En outre, la linéarité de la carte, comme je l'ai mentionné, n'est pas due à une connaissance de la morphologie des chromosomes. Le débat avec Castle en est la preuve : le choix d'une représentation unidimensionnelle ou tridimensionnelle doit, d'après les protagonistes du débat, être guidé par la recherche de l'efficacité prédictive. Quelles que soient les intuitions de chacun sur la structure réelle des chromosomes, elles ne semblent pas, à leurs yeux, devoir être décisives. Autrement dit, leur débat porte sur la recherche de *la meilleure présentation des données*, indépendamment de la question de l'explication mécanique apportée par le modèle. En conséquence, si les généticiens du groupe de Morgan utilisent ces cartes pour tirer des inférences sur la morphologie du chromosome, et non plus seulement sur les probabilités de crossing-over génétique, on peut juger qu'ils en font un usage heuristique qui excède le contenu strict des cartes, le seul qui ne soit pas sujet à controverse, et qui soit strictement *mendélien*. En tant qu'elles font partie du domaine de la génétique mendélienne, ces *cartes de liaison* semblent ne dire rien de plus que les tableaux de données numériques. En leur attribuant une autre signification – qu'on la juge légitime ou non –, Morgan et ses élèves sortent du domaine du mendélisme pur. Affirmer que la suite de l'histoire – les découvertes cytologiques de Painter (1934a) –, en leur donnant raison, change la signification des cartes reviendrait à faire de celle-ci une question de fait, ce qui semble peu raisonnable.

Pourtant, comme je vais le montrer dans la prochaine section, cette conclusion n'est pas satisfaisante et il convient de pousser l'analyse un peu plus loin. En effet, assigner à l'hypothèse mécanique qui accompagne la construction de ces cartes une fonction strictement heuristique et affirmer que ces cartes ne disent rien de plus que les tableaux de données numériques ne permet pas de rendre compte de la manière dont elles permettent de tirer des prédictions *au niveau génétique*.

6.2 Les cartes génétiques : un nouveau format

Comme je l'ai souligné au chapitre 7, l'introduction d'un nouveau format de représentation, quand bien même il serait initialement destiné à faciliter certaines inférences, joue souvent un rôle décisif dans la compréhension et le développement d'une théorie. La technique de cartographie mise au point par Sturtevant (1913a) est née de l'hypothèse mécanique de crossing-over cytologiques expliquant les crossing-over génétiques ; c'est cette hypothèse qui est à l'origine de l'adoption d'un modèle de carte *linéaire*. Quand bien même la pertinence d'une présentation graphique des données statistiques n'est pas remise en cause par le refus de cette explication au niveau chromosomique, *le choix de ce nouveau type de représentation est sous-tendu par cette hypothèse*. Introduire la notion même de carte génétique, accompagnée de celle de distance génétique, c'est aborder le problème de la transmission des facteurs héréditaires comme un problème mécanique, c'est-à-dire un problème relatif au comportement d'entités matérielles dans l'espace et le temps. C'est précisément ce que refuse Castle : avant même que des preuves cytologiques soient apportées en faveur du groupe de Morgan, on peut affirmer que les cartes génétiques ne sont pas dans le même format pour les différents protagonistes de ce débat. À supposer même qu'ils croient s'entendre sur les règles de construction des cartes, Castle et Morgan n'y voient pas la même chose, tout comme Feynman et Dyson ne voient pas la même chose dans les diagrammes de Feynman (voir chapitre 7, page 409).

Or, et c'est là le point principal de mon propos, ces différences de vues n'ont pas seulement des conséquences sur la question de savoir si les cartes représentent adéquatement la structure des chromosomes. Contrairement à ce que semblent affirmer Castle d'une part et le groupe de Morgan de l'autre, qui prétendent chacun que leur modèle est préférable *parce qu'il rend mieux compte des données* et non parce qu'il représente les mécanismes à l'œuvre au niveau chromosomique, l'hypothèse explicative favorite de chacun guide sa compréhension même de la signification *génétique* – et non pas seulement cytologique – des cartes. Ce que j'appelle la signification *génétique* des cartes, c'est l'ensemble des inférences qu'elles permettent de tirer au niveau génétique, celles dont j'ai supposé précédemment qu'elle n'étaient pas sujettes à controverse. Comme je vais le montrer, les règles de construction et d'interprétation des cartes au niveau génétique ne sont pas comprises de la même manière par Castle et par Morgan et ses élèves, en raison même de leurs divergences de vues concernant l'hypothèse explicative au niveau cytologique.

Rappelons une des critiques faites par Castle à la carte de Morgan et Bridges (1916), présentée à la figure 71 : cette carte conduit, selon Castle, à des prédictions impossibles, puisqu'elle présente des distances de plus de 50 unités cartographiques,

correspondant à des fréquences de crossing-over de plus de 50%. Or, affirme Castle, les fréquences de crossing-over ne peuvent pas dépasser 50%. Cette critique repose, comme je l'ai dit, sur un contresens de Castle relatif à la notion de crossing-over : dans sa perspective, qui consiste à refuser l'hypothèse de linéarité de l'ordonnement des gènes sur le chromosome, les doubles crossing-over n'ont pas lieu d'être, et les crossing-over *observés* (c'est-à-dire les recombinaisons effectives de facteurs) correspondent aux crossing-over *réels*. Comme il le dit, la fréquence des crossing-over observés ne peut pas dépasser 50%. De là il déduit tout naturellement que la fréquence des crossing-over réels, indiquant la distance qui sépare deux facteurs, et représentée par la distance sur la carte, ne peut pas dépasser 50%.

On peut s'étonner d'une telle confusion de la part de Castle, voire supposer qu'il fait preuve d'une certaine mauvaise foi sur ce point – qui n'est pas son seul argument contre la carte de Morgan et Bridges. Ce serait, à mon avis, une analyse trop rapide. Ce que je voudrais montrer, c'est que *la carte du groupe de Morgan est inintelligible si l'on n'admet pas qu'aux distances génétiques correspondent effectivement des distances physiques*. Autrement dit, comprendre le « mode d'emploi » de leur carte implique l'adoption préalable du modèle mécanique du crossing-over. Lors de ma présentation de la technique de Sturtevant (1913a), j'ai insisté sur les différents sens des notions de crossing-over et de distance ; j'ai montré que la distance cartographique, quand bien même elle ne correspond pas précisément à la distance physique entre les loci d'un chromosome, a malgré tout la signification d'une distance physique, puisqu'elle prétend correspondre aux crossing-over réels, et non pas aux crossing-over observés au niveau génétique (voir page 498). Si l'on refuse l'hypothèse selon laquelle les échanges de facteurs au niveau génétique s'expliquent mécaniquement par l'échange de portions de chromosomes (ou de tout autre support physique des facteurs), cette notion de distance cartographique est d'une complexité qui la rend injustifiable. Si, vraiment, on fait abstraction de l'explication mécanique suggérée par le schéma linéaire du collier de perles (figure 55), alors Castle a raison de rejeter cette notion de distance, dont la définition dépend de l'hypothèse de crossing-over multiples.

On peut, certes, s'accorder sur l'utilité de présenter graphiquement les proportions de crossing-over sans admettre l'hypothèse mécanique de Morgan ; mais on n'a aucune raison de construire les cartes de cette manière et de définir ainsi la notion de distance. Telle qu'elle est définie par le groupe de Morgan, la distance n'est pas une fonction linéaire de la fréquence de recombinaison ; les gènes séparés par une grande distance ne sont pas aussi susceptibles de faire l'objet d'une recombinaison que ne le suggère leur éloignement sur la carte. En ce qui concerne la prédiction de crossing-over, les cartes ainsi construites ne sont pas les meilleurs outils pour faciliter les inférences : les

lire correctement implique que l'on sache que les grandes distances ne correspondent pas à des probabilités de recombinaison. Savoir cela, c'est prendre en compte la possibilité de crossing-over multiples et d'effets d'interférence, *ce qui fait intervenir une hypothèse mécanique*.⁵¹

En conséquence, la distance sur la carte joue un rôle différent de la simple mise en diagramme des données numériques, au sens où *il convient d'avoir en tête le modèle mécanique (schématique) du collier de perles pour comprendre* qu'une plus grande distance n'implique pas, proportionnellement, une plus grande probabilité de recombinaison. Ainsi, la spatialisation des données n'est pas simplement destinée à aider l'utilisateur à tirer des inférences plus rapidement qu'il ne le peut au moyen d'un tableau présentant les mêmes données. La forme linéaire de la carte impose des propriétés particulières aux relations entre les pourcentages de recombinaison, et donc la croyance que ces propriétés sont respectées dans la nature.⁵² Je suggère donc que l'usage du mot « distance », défini initialement comme le pourcentage de recombinaison effective, puis comme le pourcentage de crossing-over *réels*, n'est pas seulement métaphorique. La distance sur la carte doit bien permettre à l'utilisateur de visualiser une distance physique. Ainsi, le caractère explicatif de l'hypothèse mécanique du collier de perles joue un rôle crucial dans la lecture de la carte. Sans cette hypothèse, on ne peut pas comprendre et utiliser correctement la carte pour tirer des prédictions, même au niveau génétique. En ce sens, le modèle du collier de perles n'est pas strictement analogique : si l'on n'admet pas que, *réellement*, les gènes sont des particules matérielles ordonnées linéairement, on ne peut pas vraiment comprendre les cartes du groupe de Morgan.

Insistons sur un point important : cela ne veut pas dire que l'adoption de l'hypothèse de linéarité implique l'adoption de la théorie chromosomique – on a vu que Bateson admet la première en refusant la seconde – mais qu'elle implique la croyance en une disposition linéaire des gènes, quel qu'en soit le support physique. En ce sens, les cartes génétiques restent du domaine du mendélisme pur, mais *contiennent une hypothèse mécanique qui ne figure pas dans les tableaux de données numériques* en tant que tels. C'est en ce sens que je propose de les considérer comme *schématiques*, sans pour autant affirmer qu'elles sont des schémas *de chromosomes*. Notons que cette conclusion fait écho à une remarque que je fais, au chapitre 7 (section 1.3.4, page 385), à propos de la relation entre diagrammes, schémas, et représentations linguistiques. J'y suggère en effet que les diagrammes sont « plus proches » des représentations linguistiques que les schémas, au sens où la mise en diagramme (ou en

⁵¹La possibilité de crossing-over multiples est pour ainsi dire l'hypothèse nulle si l'on adopte un modèle de carte linéaire, mais devient effectivement une hypothèse *ad hoc* dans le modèle de Castle.

⁵²Merci à Patrice David pour cette formulation.

graphe) n'ajoute généralement pas d'hypothèse aux énoncés auxquels le diagramme (ou le graphe) correspond (en revanche, elle facilite certaines inférences, comme on l'a vu). Affirmer que les cartes génétiques sont des diagrammes reviendrait à dire qu'elles ne contiennent rien de plus que les données numériques. Défendre l'idée selon laquelle ce sont des schémas – et donc affirmer que les relations spatiales y tiennent lieu de relations spatiales – permet de rendre compte du fait que l'introduction du format cartographique par Sturtevant et son utilisation par le groupe de Morgan implique bien l'adoption d'une hypothèse nouvelle. Cette hypothèse – celle de la linéarité, qui sous-tend le modèle mécanique du crossing-over – n'est donc pas seulement une analogie suggestive : elle est proprement *indissociable* de l'utilisation des cartes.

On a vu que l'adoption de l'hypothèse mécanique du crossing-over est, historiquement, due à l'adoption de l'hypothèse chromosomique. J'ai montré, cependant, que ces deux hypothèses (génétique et cytologique) sont conceptuellement indépendantes. Les cartes génétiques appartiennent bien au domaine du mendélisme pur. Cependant, l'adoption du modèle mécanique du crossing-over constitue en elle-même une avancée conceptuelle au sein du mendélisme qui lui fait faire un pas considérable en direction de l'hypothèse chromosomique, en offrant une compréhension des phénomènes de recombinaison comme le résultat d'un mécanisme *spatial*.

6.3 Versions du mendélisme : différents « modes d'emploi » des représentations théoriques

J'ai suggéré à plusieurs reprises que l'analyse de ce que j'appelle le « symbolisme mendélien » permet d'isoler un mendélisme « pur », conceptuellement indépendant des hypothèses cytologiques destinées à expliquer les phénomènes génétiques. En tant que présentations graphiques de données statistiques mendéliennes, les cartes génétiques font authentiquement partie de la génétique mendélienne au sens strict ; de nombreux mendéliens non-morganiens (Castle, Bateson) acceptent le principe de la cartographie en refusant le modèle du collier de perles (qui inclut l'hypothèse chromosomique, refusée par Bateson, *et* l'hypothèse de la linéarité, refusée par Castle). Cependant, comme je viens de le montrer, selon que l'on accepte ou non ce modèle explicatif, on ne lit pas les cartes de la même manière, et on n'y voit pas la même chose. Dans la mesure où les règles de construction des cartes et la définition de la notion de distance par Morgan et son groupe sont sous-tendues par l'hypothèse chromosomique, on peut conclure que leur compréhension et leur pratique du mendélisme sont elles-mêmes déterminées par leur engagement en faveur de l'hypothèse

chromosomique.⁵³

En conséquence, isoler le mendélisme par l'analyse du type de représentations qui le caractérise permet de mettre en évidence que, en pratique, les engagements des théoriciens en faveur de différents types d'hypothèses – qui ne font pas, en tant que telles, parties du mendélisme pur – interagissent avec la manière dont ils comprennent et développent le mendélisme. Autrement dit, après avoir distingué le niveau théorique du mendélisme du niveau théorique de la cytologique, on est conduit à constater que, dans la pratique, ces différentes théories n'existent pas à l'état « pur ».

Je suggère que ces différentes façons de comprendre et de pratiquer la génétique mendélienne, qui se manifestent dans la manière dont chacun construit, utilise et interprète différents types de représentation, peuvent être décrites comme des *versions* du mendélisme, au sens où je parlais, au chapitre 3, de « versions » de la mécanique classique. Le groupe de Morgan⁵⁴, Castle et Bateson ont chacun, en ce sens, leur version du mendélisme. Ces versions, que j'ai proposé de définir comme l'ensemble des représentations mentales associées à la pratique d'une théorie, sont le résultat de leur formation, de leurs engagements et de leurs préférences théoriques, explicites ou non. Si ces différentes composantes sont inaccessibles à l'historien, il est possible, comme je viens de le faire, de caractériser leurs versions par leurs propos explicites, ainsi que par ce que l'on peut bien appeler leurs « modes d'emploi » de différents types de représentation : ces pratiques représentationnelles ont en effet le grand avantage de rendre partiellement accessibles à l'analyste les raisonnements menés par les différents théoriciens, sans qu'il soit nécessaire d'avancer des hypothèses invérifiables sur leurs engagements implicites.

L'analyse de ces différentes versions de la génétique mendélienne semble corroborer l'hypothèse du « réalisme sélectif », avancée par Humphreys (2004) et exposée au chapitre 6 (section 3.2.4). On vient de voir, en effet, que l'utilisation d'une même représentation peut faire l'objet de différents types d'engagements de la part des agents : tout en s'accordant – partiellement – sur certaines règles de construction des cartes,

⁵³ Il en va certainement de même pour les préférences supposées de Castle pour des modèles chimiques ; il est cependant plus difficile d'analyser le lien entre ces engagements et son modèle de carte, dans la mesure où il ne propose pas de modèle explicatif clair et détaillé. Il n'y a cependant pas lieu, selon moi, de distinguer nettement, comme le fait Wimsatt (1987), entre un Castle instrumentaliste et un Morgan réaliste. Cette histoire n'est pas tant l'exemple de la victoire d'un programme réaliste sur un programme instrumentaliste, mais, assez trivialement, celle d'une hypothèse vraie sur une hypothèse fausse.

⁵⁴ J'ai considéré, au cours de cette analyse, que le groupe de Morgan entretenait une version unique du mendélisme. On pourrait envisager de mener une enquête sur les différentes versions entretenues par chacun de ses membres, sur la base de leur pratique et des directions dans lesquelles chacun a conduit les développements de la génétique.

leurs utilisateurs peuvent être en désaccord sur le statut référentiel de leurs différents éléments. Qui plus est, les débats explicites sur ces questions conduisent chacun à préciser et à développer sa propre interprétation des cartes. Ce travail est, comme je l'ai suggéré au chapitre 7, constitutif de l'activité théorique, dans sa double dimension représentationnelle et computationnelle. Les préoccupations relatives à l'utilité prédictive des cartes sont en effet indissociables des engagements théoriques qu'elles suggèrent – et qui en sous-tendent la construction.

Pour finir, je souhaite tirer de cette étude de cas quelques considérations sur l'histoire des sciences. Au risque d'être un peu caricaturale, disons que l'on oppose traditionnellement l'histoire conceptuelle – réservée aux philosophes – et l'histoire des pratiques – réservée aux historiens. J'espère être parvenue à montrer, dans ce qui précède, que cette distinction mérite d'être nuancée. L'examen que j'ai proposé des débats sur la cartographie génétique vise à élucider la nature des différentes hypothèses qui constituent la génétique classique. C'est à la manière dont ses différentes composantes théoriques s'articulent que je me suis intéressée ; en ce sens, si l'on peut tirer de mon examen quelques éléments pour une histoire de la génétique classique, ils s'apparentent incontestablement à une histoire conceptuelle et non à une histoire du contexte sociologique et institutionnel de la génétique des années 1920. Pour autant, c'est par une attention aux individus et à leur pratique de la génétique que j'ai cherché à éclairer l'articulation des composantes théoriques de la génétique classique. La notion de version, en tant qu'elle désigne la manière dont un agent, formé dans un contexte particulier et habitué à manipuler un certain type de représentation, comprend les hypothèses théoriques qu'il apprend et développe, vise ainsi à rendre compte du caractère situé et pratique de l'activité théorique, en tant qu'activité conceptuelle. Autrement dit, ce que j'espère être parvenue à montrer, c'est que l'histoire conceptuelle peut et doit aussi être une histoire des pratiques.

Conclusion

Dans ce chapitre, j'espère avoir montré la fécondité d'une analyse de l'activité théorique centrée sur l'utilisation de certains types de représentations par les agents. Ces représentations – en l'occurrence, les cartes génétiques – ne sont pas tant le moyen d'accéder à « la théorie » dont elles seraient l'expression, que l'objet même de l'activité théorique.

La construction et l'interprétation des cartes par les généticiens est, en ce sens, le lieu même de l'élaboration de leur version de la génétique. Par contraste avec une analyse des relations inter-théoriques fondée sur la comparaison de schémas explicatifs hypothético-déductifs, l'examen de la manière dont les types de représentation

associés à différents cadres théoriques (mendélisme et cytologie) s'articulent dans la pratique des généticiens permet de mettre en évidence la possibilité, au sein de ce que Kuhn appellerait une même « matrice disciplinaire », de différentes versions.

Ces versions, qui se manifestent dans le cas étudié par des débats explicites et quelquefois des malentendus, sont, comme on l'a vu, liées aux engagements théoriques de chacun. Dans la mesure où ces engagements sont parfois implicites, un bon moyen de les analyser est ainsi d'étudier leurs conséquences sur les différentes manières d'utiliser les représentations théoriques.

L'alternative, comme on l'a vu, n'est pas aussi simple que celle de l'acceptation ou du rejet de l'hypothèse chromosomique : l'analyse de l'interprétation des cartes par chacun des généticiens marque la possibilité de comprendre cette hypothèse elle-même de différentes manières. Il en résulte, comme dans le cas des diagrammes de Feynman étudié au chapitre 7, la possibilité pour différents agents d'utiliser les « mêmes » représentations en s'accordant sur certaines règles de construction, tout en y « voyant » des choses différentes.

Ces considérations permettent, comme j'espère être parvenue à le montrer, de saisir un aspect essentiel *du théorique*, en tant qu'il qualifie un ensemble d'activités consistant à représenter les phénomènes au moyen d'hypothèses employant des concepts dont les référents, qu'il s'agisse des forces ou de l'énergie en mécanique classique, ou, dans le cas présent, des gènes, n'ont pas de statut ontologique clairement défini⁵⁵ (sans quoi ils ne seraient plus *théoriques*). Dans la mesure où l'activité théorique consiste à prédire et expliquer les phénomènes en les représentant au moyen de ces concepts, la forme même des représentations, et la manière dont elles sont utilisées, joue un rôle central.

Pour terminer, je voudrais suggérer qu'une telle analyse de la génétique classique pourrait permettre de jeter un éclairage nouveau sur certaines questions relatives à l'histoire de cette science. Comme je l'ai noté dans l'introduction du chapitre 8, l'histoire de la génétique est souvent décrite comme celle de la *matérialisation progressive des gènes* (voir Jacob, 1970; Gayon, 2000). Le rapprochement du symbolisme mendélien abstrait avec les représentations schématiques et picturales des chromosomes manifeste cette évolution. Par l'analyse qui précède, j'espère être parvenue à mettre en évidence l'articulation complexe, dans l'utilisation des cartes génétiques, des différentes composantes théoriques dont elles marquent la rencontre, dans une perspective *synchronique* et non pas seulement *diachronique*. On pourrait finalement envisager, par une extension de mon analyse à celle de l'évolution diachronique des représentations en génétique, de déceler, à chaque étape, différentes versions du mendélisme

⁵⁵ Je veux dire par là que ni leur nature ni même leur existence ne sont clairement établies.

et, ce faisant, d'éclairer certaines questions concernant les relations entre génétique classique et génétique moléculaire.⁵⁶

⁵⁶Selon Patrice David (communication personnelle), une telle perspective permettrait de montrer que la plupart des biologistes actuels, qui ont souvent l'impression de pratiquer une science unifiée, ont pourtant un « cerveau dissocié » qui a recours selon les moments à des versions différentes de la génétique ; leur capacité à passer de l'une à l'autre en « enjambant un abîme », pour reprendre la formule de Patrice David, est sans doute le résultat d'un « long entraînement ».

Conclusion de la troisième partie

Cette troisième et dernière partie visait deux objectifs principaux. Le premier consistait à montrer la fécondité de la méthode et des outils proposés dans les parties précédentes pour une analyse de l'activité théorique, en tant qu'elle ne se réduit pas à l'application de théories pouvant être reconstruites sous la forme d'un ensemble hypothético-déductif d'énoncés. Le second était d'éclairer, par ce moyen, la nature de ce que l'on appelle « mendélisme », et l'articulation de ce dernier avec la théorie chromosomique de l'hérédité.

Le choix d'un domaine qui n'appartient pas aux sciences physiques, sur lesquelles s'est longtemps concentrée l'attention des philosophes des sciences, reposait sur la conviction selon laquelle il y a quelque chose de commun à toutes les activités proprement théoriques. Ces dernières consistent, en grande partie, à utiliser, pour prédire et expliquer certains phénomènes, des représentations faisant référence à des entités, des structures, ou des processus, dont la nature et l'existence peuvent faire l'objet de différents types de croyances et d'engagements.

La construction et l'utilisation de ces représentations, dans la mesure où elles doivent aussi servir d'outils d'inférence, sont à la fois subordonnées à des impératifs prédictifs et guidées par les engagements théoriques de leurs utilisateurs. L'analyse de l'utilisation des cartes génétiques est un exemple d'une analyse des pratiques des agents en tant qu'elles sont éminemment théoriques.

J'ai ainsi, au chapitre 8, isolé ce que j'ai appelé le « mendélisme pur », en montrant qu'il consiste en l'utilisation d'un certain type de représentation. Cela m'a permis de le distinguer d'une autre composante de la génétique classique, celle de l'étude des chromosomes. En comparant cette approche avec celle de Kitcher (1984), j'ai suggéré qu'elle permettait d'aborder de façon nouvelle la question des relations inter-théoriques.

Le chapitre 9 a mis cette suggestion en œuvre. Cela m'a permis, par l'analyse de la construction et de l'utilisation des cartes génétiques, de généraliser la notion de version, proposée au cours de l'examen de la mécanique classique, et d'analyser le jeu des engagements théoriques dans la pratique des généticiens et leur compréhension du

mendélisme. Dans ma perspective, les versions de la génétique ne correspondent pas tant, comme chez Kitcher, à la mise en œuvre de différents schèmes de raisonnements pour résoudre des problèmes de pedigree, qu'à différentes manières de construire et de comprendre les cartes génétiques, et, ce faisant, de se représenter les phénomènes de l'hérédité.

Conclusion générale : les modes d'emploi de la théorie



Cornelius Troost, 1741, *Les Mathématiciens ou la jeune fille qui s'enfuit : la Querelle entre les docteurs Raasbollius et Urinaal*, Mauritshuis, La Haye.

Ma démarche est fondée sur l'adoption d'un parti pris méthodologique, qui consiste à aborder l'étude de l'activité théorique du point de vue des sujets de cette activité. J'espère être parvenue, au terme de ce travail, à montrer la pertinence et la fécondité de cette méthode. J'espère aussi, par les analyses qu'elle m'a permis de mener, avoir donné quelques arguments convaincants en faveur de la thèse principale que j'ai cherché à défendre : une théorie scientifique, dès lors qu'on cherche à en définir le contenu sans prendre en compte les individus, experts ou profanes, qui l'utilisent et la comprennent, n'est tout simplement plus une théorie, si l'on entend par là un ensemble d'hypothèses permettant de représenter, de prédire et d'expliquer les phénomènes. Pour conclure ce travail, je me propose de revenir sur les principaux arguments qui m'ont permis de défendre cette thèse, et d'esquisser quelques questions que ces arguments soulèvent.

1 Que reste-t-il de la théorie ?

J'ai affirmé, à plusieurs reprises, que la notion de théorie, telle qu'elle est souvent conçue en philosophie des sciences, n'est pas une unité d'analyse bien adaptée à l'étude de l'activité théorique. J'ai apporté des arguments de différents types en faveur de cette thèse. Dans cette section, je les rappelle brièvement, afin de préciser la manière dont, au terme de ce travail, je propose de concevoir la notion de théorie.

Un premier ensemble d'arguments m'a été fourni par le cas de la mécanique classique. L'examen de cet exemple canonique de théorie scientifique m'a permis de mettre au jour un certain nombre de difficultés qui m'ont conduite à réexaminer les notions de contenu et d'équivalence au nom desquelles on affirme habituellement l'identité de cette théorie. D'une part, les différentes formulations sous lesquelles elle est aujourd'hui enseignée et utilisée ne représentent pas les phénomènes de la même manière et ne permettent pas les mêmes inférences. S'il existe quelque chose comme le contenu de la mécanique classique (par delà ses différentes formulations), son analyse ne permet pas de rendre compte de la manière dont les équations de la mécanique sont utilisées, en pratique, pour expliquer et prédire les phénomènes (chapitre 2). D'autre part, les différentes tentatives qui ont été faites, au cours de l'histoire, pour présenter le contenu objectif de la mécanique sous la forme d'un ensemble déductif de principes ont révélé les limites que rencontre inévitablement une telle entreprise. L'examen de certaines de ces présentations, que j'ai proposé de considérer comme l'expression d'autant de versions de la mécanique, m'a permis de remettre en question la possibilité d'une distinction nette entre différence formelle et différence conceptuelle. Selon la manière dont on organise les différents principes de la mécanique, et selon la manière dont on les rattache aux phénomènes empiriques,

la mécanique – et ses différents principes et concepts – n’ont pas exactement la même signification (chapitre 3). Ce premier ensemble d’arguments vise donc à montrer, par l’exemple, qu’une théorie comme la mécanique classique, en tant qu’elle est toujours présentée sous une forme et depuis une perspective particulières, ne peut pas être considérée comme une unité monolithique indépendante des esprits individuels qui la comprennent.

En raison même du caractère canonique de la mécanique, cette analyse suffit à elle seule à remettre sérieusement en question la notion de théorie au nom de laquelle on affirme généralement que la mécanique est bien une seule et même théorie. J’ai cependant cherché à apporter un deuxième type d’arguments en faveur de ma démarche. Ces arguments, à la différence de ceux que j’ai proposés dans la première partie de mon travail, visent à fonder cette démarche en montrant les limites internes des entreprises qui ignorent, dans l’analyse du contenu des théories, la place de leurs utilisateurs. L’argument du chapitre 5 cherche ainsi à montrer que certaines tensions internes au projet formaliste en philosophie des sciences sont dues au fait qu’il y a quelque chose, dans la notion de théorie elle-même – conçue comme outil de représentation et d’inférence –, qui résiste aux analyses fondées sur l’idée d’une distinction nette entre forme et contenu. J’ai ensuite examiné les thèses de certains philosophes qui critiquent ces approches formelles, et dont mon travail est l’héritier. Par cet examen, j’ai cherché à montrer que certaines limites (et parfois contradictions internes) de leur entreprise sont dues, à leur tour, au maintien d’une conception des théories encore trop abstraite et éloignée de la pratique : l’analyse critique des propos de Kuhn sur l’incommensurabilité (au chapitre 6), mais aussi celle de la manière dont Kitcher envisage ce qu’il appelle la « pratique » et les « versions » de la génétique classique (au chapitre 8) m’ont ainsi permis de préciser et de justifier ma démarche et mes outils.

Que reste-t-il, au terme de ces analyses, de la notion de théorie ? Comme j’ai cherché à le montrer dans ma seconde étude de cas, sur la génétique classique, le but principal que je poursuis n’est certainement pas de « réduire » la théorie à la pratique, ni d’affirmer que les produits de l’activité scientifique ne sont *que* les traces⁵⁷ concrètes construites par les scientifiques au cours de leur travail. Affirmer que la notion classique de théorie doit être décomposée, et qu’il convient d’analyser le détail de la représentation et de l’inférence scientifiques, c’est dire que le contenu de la connaissance scientifique est avant tout le fruit d’une *activité*, dont le lieu même est l’interaction cognitive des agents avec les représentations qu’ils utilisent. Le produit de cette activité, les théories, conçues comme des ensembles d’hypothèses, jouent un

⁵⁷ Voir par exemple les travaux de Latour et Woolgar (1979).

rôle fondamental dans la vie cognitive, et plus particulièrement dans la compréhension scientifique ; l'erreur, selon moi – si l'on cherche à rendre compte de la manière dont ces hypothèses permettent de représenter et de prédire les phénomènes – consiste à les rigidifier et à les concevoir indépendamment des esprits qui les produisent et les comprennent. Étudier ces hypothèses en prêtant attention aux représentations concrètes qui les expriment et servent à la prédiction et à l'explication des phénomènes permet, comme j'espère l'avoir montré par mon examen de la technique de cartographie génétique (chapitre 9) d'éclairer la nature, le statut et l'interaction de ces différentes hypothèses.

C'est afin de rendre compte de la dimension cognitive de cette activité que j'ai proposé la notion de version. La version d'un agent, c'est, comme je l'ai souligné au début de ce travail, la théorie pour l'agent, la théorie en tant qu'elle est comprise et utilisée par cet agent. Selon la manière dont il relie ces différentes hypothèses entre elles et au monde empirique, elles n'ont pas exactement la même signification. L'analyse de la notion de format (chapitre 7) m'a permis de préciser et de renforcer la thèse selon laquelle la signification d'une représentation théorique est indissociable de la manière dont les agents l'utilisent dans leurs raisonnements. Ces raisonnements sont eux-mêmes doublement déterminés par la forme de la représentation et par les capacités cognitives de l'agent, ainsi que par ses engagements et ses intérêts théoriques. Au cours de mon analyse des versions de la mécanique, tout comme lors de mon examen de la notion de format, prendre au sérieux les différences de *forme* (si l'on entend par là aussi bien les différences architectoniques entre les présentations de la mécanique que les différents types de représentations particulières – énoncés, équations, diagrammes, etc.), et les considérer comme des différences conceptuelles, m'a conduit, à chaque fois, à m'intéresser aux différences cognitives inter-individuelles. Autrement dit, de considérations sur la forme (objective) des représentations, j'ai été conduit à souligner l'importance des particularités cognitives des utilisateurs. Ce mouvement – de la différence de forme à la différence cognitive – soulève une question difficile, que je souhaite brièvement énoncer.

2 Différences épistémiques et différences psychologiques : quelle frontière ?

Une des motivations initiales de ce travail est le constat de l'importance des différences de forme entre des représentations théoriques (ou des ensembles de représentations théoriques) jugé(e)s équivalent(e)s par les approches classiques de la connaissance scientifique. Ces différences, dans le cas de la mécanique, concernent la

forme des équations utilisées pour prédire et expliquer le mouvement des corps, ainsi que l'organisation déductive des principes et des concepts de cette théorie. C'est l'examen de ces différences qui a motivé mon refus d'une étude de la science *en principe*, et l'adoption d'une conception du contenu des représentations qui prend en compte la manière dont, *en pratique*, ces représentations sont utilisées pour tirer des inférences. Selon cette dernière conception, le contenu d'une équation comme la deuxième loi de Newton n'est pas défini par tout ce qui peut en être déduit en principe, en vertu des règles de la logique et des mathématiques, mais par ce qu'il est effectivement possible, en pratique, d'en tirer, et par les processus inférentiels que cela implique de mettre en œuvre.

Autrement dit, à la différence de la perspective *en principe* qui s'accompagne d'une conception logico-empirique de la signification, j'ai cherché à montrer que la différence formelle entre les équations de Newton et celles de Lagrange implique une différence de contenu : parce que ces équations *ne permettent pas d'accéder* aux mêmes informations – parce que leur différence formelle implique en ce sens une différence *épistémique* – leur contenu *en pratique* n'est pas le même. Cette différence, bien qu'elle soit gommée par une approche logico-empirique du contenu, est toutefois objective : il n'est pas nécessaire de faire appel aux particularités cognitives des agents pour montrer la différence entre une équation de Newton et une équation de Lagrange – ou entre une liste de données numériques et un tableau – ; c'est parce qu'elles ont une forme différente que l'on ne peut pas en déduire, en pratique, la même chose. De même, les différences architectoniques entre la mécanique fondée sur un principe différentiel (comme la deuxième loi de Newton) et la mécanique fondée sur un principe variationnel (comme le principe de Hamilton) sont objectives : l'ordre déductif des principes n'est objectivement pas le même. Les particularités cognitives des agents ne semblent pas devoir entrer en jeu dans l'exposé de ces différences. Ce n'est, certes, que pour des agents aux capacités cognitives limitées que ces différences formelles impliquent des différences de contenu, mais elles semblent de prime abord être les mêmes pour *tous les agents cognitifs*. La différence épistémique entre les équations de Newton et les équations de Lagrange semble bien être la même pour tous. Si le « mode d'emploi » des équations de Newton et des équations de Lagrange n'est pas le même, c'est en raison de leur différence de forme, qui n'appelle pas les mêmes inférences, et non des particularités cognitives des différents agents. Autrement dit, il semble exister une frontière nette entre une différence épistémique – dont on peut montrer qu'elle est due à une différence de forme – et une différence psychologique.

Cependant, en cherchant à rendre compte de ces différences de forme comme d'authentiques différences de contenu, aussi bien lors de mon analyse de l'exemple de la mécanique que lors de mon examen de la notion de format (destinée à caractériser la

différence entre, par exemple, une liste de données numériques et un tableau), j'ai été conduite à prendre en compte les particularités cognitives des agents individuels. Ce « glissement » dans mon analyse reflète le cheminement qui a été le mien au cours de ce travail. En cherchant d'abord à exposer clairement les différences objectives entre les formulations de la mécanique – différences qui me semblaient, en tant que telles, remettre en cause les conceptions logico-empiriques du contenu des théories – je me suis aperçue que les différents savants qui s'y sont intéressés (ceux dont j'ai présenté les versions au chapitre 3) ne décrivent pas ces différences de la même manière. Il m'a alors semblé inévitable de prendre les variations inter-individuelles au sérieux ; parce qu'une différence de forme n'implique une différence de contenu que pour des agents cognitifs, si ces différences de forme n'ont pas les mêmes conséquences pour différents agents cognitifs, alors on doit aussi considérer les différences inter-individuelles pour caractériser le contenu d'une théorie. C'est ainsi que j'en suis arrivée à proposer l'hypothèse des versions.

C'est un mouvement analogue que j'ai suivi au cours de mon analyse de la notion de format, présentée au chapitre 7. J'ai, d'abord, voulu montrer que des représentations de différents types – formules et figures – ne jouent pas le même rôle inférentiel pour des agents aux capacités cognitives limitées. Ces différences, que j'ai appelées des différences de *format*, me semblaient d'abord pouvoir être caractérisées objectivement, en faisant référence aux propriétés des systèmes symboliques dans lesquels elles fonctionnent. Cependant, ici encore, la prise en compte des processus inférentiels que la manipulation de différentes représentations exige de la part des agents m'a conduite à constater que la frontière entre une différence épistémique objective, valant pour tous les agents, et une différence cognitive ou psychologique, en vertu de laquelle deux agents n'emploient pas la même représentation de la même manière, ne peut pas être fixée de manière précise. Autrement dit, constater que deux représentations « équivalentes » n'ont pas le même mode d'emploi conduit naturellement à reconnaître qu'il y a, pour ainsi dire, *autant de modes d'emploi* que d'agents.

Tirant les conséquences de ce constat, je me suis aperçue que l'hypothèse des versions, ainsi que l'idée selon laquelle chacun, selon ses capacités cognitives, ses engagements et ses intérêts théoriques, ne fait pas précisément le même usage des représentations théoriques dans ses raisonnements et sa pratique, permettent de rendre compte de plusieurs aspects importants de l'activité théorique. Comme je l'ai plusieurs fois souligné, adopter cette perspective cognitive permet de rendre compte de ce que les processus d'apprentissage et de recherche théoriques ont en commun. L'étudiant qui apprend une théorie, mais aussi l'expert qui la développe, approfondissent leur compréhension de cette théorie en apprenant ou en inventant de nouveaux chemins inférentiels entre ses différentes hypothèses.

À ce point, se pose naturellement la question de la *limite* qu'une analyse de la connaissance scientifique peut assigner entre ce qui relève d'une authentique différence épistémique et ce qui relève d'une différence psychologique. Plus précisément, la question est de savoir quelles différences psychologiques doivent être traitées *comme* des différences épistémiques, et donc être prises en compte par une analyse de la connaissance scientifique.⁵⁸ Il va de soi que, pour une étude du contenu de la mécanique classique, les différences de forme entre les équations de Newton et de Lagrange sont plus intéressantes que les différences cognitives entre un profane ignorant du calcul différentiel et un expert. Cependant, quoique la différence soit immense, j'en suis arrivée à la conclusion que c'est une différence de degré, et non de nature.

Selon ce que l'on cherche à étudier, la limite entre différences épistémiques et différences psychologiques varie. Pour étudier les progrès effectués par Hamilton, les différences de forme mathématique sont sans doute les seules qu'il convienne de prendre compte. Pour étudier certaines controverses scientifiques – comme par exemple les débats du dix-neuvième siècle au sujet des principes de la mécanique (chapitre 3), ou la « scission entre Feynman et Dyson » (chapitre 7) –, on a intérêt à déplacer la frontière, et à prendre en compte les engagements philosophiques et théoriques, ainsi que les domaines de spécialité de chacun (le mathématicien et le physiologiste sont susceptibles de ne pas comprendre la mécanique de la même manière). Enfin, pour étudier les processus d'apprentissage – qui, je me suis efforcée de le montrer tout au long de ce travail, présentent un intérêt authentiquement épistémologique – considérer certaines différences psychologiques comme des différences épistémiques semble inévitable.

Aborder ainsi le contenu des représentations scientifiques du point de vue des agents, et admettre que la distinction entre différence épistémique et différence psychologique est de degré, semble faire courir le risque d'une relativisation complète du contenu de ces représentations aux agents qui les utilisent ; autrement dit, dé-

⁵⁸ Précisons un point important : comme je l'ai souligné à plusieurs reprises, ce n'est pas aux particularités psychologiques des agents en tant que telles que je m'intéresse, mais à leur manifestation dans l'usage que les agents font de certaines représentations théoriques. En affirmant que les différences individuelles déterminent différentes compréhensions d'une théorie, je n'ai pas cherché à *expliquer* la manière dont les particularités cognitives de chacun influencent son utilisation d'un ensemble d'hypothèses. Ce qui m'a intéressée, tout au long de ce travail, c'est la manière dont la version de chacun se manifeste dans sa pratique et, le cas échéant, dans ses jugements explicites sur la signification des représentations théoriques. Les raisons pour lesquelles Castle, par exemple, ne croit pas que les gènes soient ordonnées linéairement sur les chromosomes n'ont pas concentré mon attention. Ce que j'ai cherché à montrer, c'est que les engagements théoriques des différents généticiens – quelle qu'en soit la cause « psychologique » – déterminent leur compréhension même du mendélisme.

crire les différences psychologiques comme des différences épistémiques peut sembler conduire à une position psychologiste, selon laquelle le contenu d'une représentation est entièrement défini par la manière dont un agent la comprend. Les analyses que j'ai conduites tout au long de ce travail dissipent, je l'espère, cette impression ; je me propose, dans la section suivante, de revenir sur ce problème.

3 Compréhension et contenu

En abordant l'étude du contenu des connaissances scientifiques du point de vue de la compréhension qu'en ont les agents, j'ai cherché à montrer que les représentations théoriques n'ont véritablement de signification qu'en tant qu'elles sont comprises et utilisées. Souligner, en plus des différences objectives de forme, les différences inter-individuelles, revient donc à affirmer que le contenu d'une représentation théorique n'est pas exactement le même pour deux agents qui ne l'utilisent pas de la même manière. Qu'est-ce qui évite de conclure, de ces considérations, à la dissolution du contenu de la connaissance théorique, et à sa relativisation complète ?

Comme je me suis efforcée de le montrer tout au long de ce travail, la compréhension d'une théorie repose sur une double compétence, que j'ai appelée représentationnelle et computationnelle. Elle consiste en la capacité à manipuler et à raisonner au moyen de représentations particulières, et à les utiliser pour résoudre des problèmes. Ces raisonnements sont extrêmement contraints par les règles objectives qu'imposent les relations logiques entre les concepts d'une théorie. En outre, l'application de ces représentations aux phénomènes empiriques suit elle aussi des procédures rigoureuses ; elle se heurte, pour ainsi dire, à la réalité empirique. Affirmer que la signification d'une représentation dépend de la compréhension qu'en ont les agents ne revient naturellement pas à dire que sa vérité ou sa fausseté dépendent du bon vouloir des agents. Ce que je me refuse à dire, ce n'est pas que les équations de Maxwell sont vraies (ou fausses), mais c'est plutôt qu'il puisse exister une procédure de « traduction » triviale qui permette de donner un sens à l'affirmation de Weinberg selon laquelle « si jamais nous découvrons des créatures intelligentes sur quelque planète lointaine et traduisions leurs travaux scientifiques, nous découvririons qu'eux et nous avons découvert les mêmes lois. » (Weinberg, 1996, p. 14, cité par Hacking, 1999, p. 107)

Renoncer, comme je l'ai fait, à définir le contenu des théories en termes vériconditionnels – ou logico-empiriques –, ne revient pas à nier qu'il existe des règles en vertu desquelles certaines inférences sont correctes et d'autres fautives, ni que les conséquences empiriques de ces inférences puissent être infirmées ou confirmées par la réalité empirique. En revanche, comme le problème que Hempel appelle le « dilemme du théoricien » le montre, c'est bien parce que leurs référents ne sont pas aussi

aisément identifiables que ceux des mots « chaise » ou « table » que les représentations des sciences empiriques sont des représentations théoriques. La conception que je défends du contenu des représentations théoriques vise à montrer que, pour ainsi dire, le dilemme du théoricien ne se pose pas : c'est parce que les règles objectives sous-déterminent l'usage que l'on peut faire de ces représentations qu'il est toujours possible d'en tirer de nouvelles conséquences.

Comme la critique que j'ai faite des propos de Kuhn sur l'incommensurabilité vise à le montrer, l'affirmation relativiste selon laquelle « le monde change » à chaque changement de paradigme ne découle pas d'une attention à la pratique et aux processus inférentiels particuliers mis en œuvre par les agents, mais bien au contraire de l'idée selon laquelle la signification d'un concept est *entièrement* déterminée par ses relations logiques avec les autres concepts de la théorie à laquelle il appartient. Comme on l'a vu, différents agents peuvent s'entendre sur certaines règles d'utilisation des représentations théoriques sans pour autant les comprendre exactement de la même manière. Prêter attention à ces différentes versions permet bien souvent d'éclairer le processus même de découverte scientifique et de rendre compte d'un authentique progrès. Ainsi, l'analyse de l'utilisation et de l'interprétation des cartes génétiques par Castle d'une part et par le groupe de Morgan d'autre part m'a permis de mettre en lumière la manière dont la version de chacun, guidée par ses engagements théoriques, le conduit à employer différemment les cartes génétiques. Cela n'empêche pas de dire, *a posteriori*, que Morgan avait raison de se représenter l'arrangement des gènes de manière linéaire, et Castle tort. Ce que j'ai cherché à souligner tout au long de ce travail, c'est finalement le constat que le contenu des connaissances théoriques est avant tout le produit d'esprits au travail. Que le contenu des hypothèses théoriques varie en fonction de la compréhension que chacun en a ne revient pas à nier qu'il y a bien « quelque chose à comprendre », et que ce quelque chose est ce qui guide l'enquête scientifique elle-même.

Annexes

Les développements qui suivent sont des annexes de la première partie de ce travail, sur la mécanique classique. Certaines d'entre elles contiennent des démonstrations qui auraient inutilement chargé le texte principal du chapitre 2. Elles sont la plupart du temps tirées des ouvrages de Lanczos (1970) et Goldstein (1950/2002). D'autres comportent des développements importants pour la mécanique classique, qui doivent faire l'objet de toute analyse approfondie de cette théorie ; je ne les ai pas intégrés au texte principal, dans la mesure où ils ne servent pas directement mon propos. Toutes ces annexes font l'objet d'un renvoi en note à différents moments du texte principal. Afin de ne pas les confondre avec celles du chapitre 2, la numérotation des équations est suivie d'un « a ».

A.1 Les mouvements non rectilignes. Moment de force et moment cinétique. Théorème du moment cinétique.

Lorsque l'on sort de l'étude des mouvements rectilignes, comme par exemple dans le cas de la description du mouvement d'un pendule, on emploie les équivalents angulaires de la position, de la vitesse et de l'accélération linéaires, qui sont aussi des quantités vectorielles.

On définit alors le *moment de force* \mathbf{N} par rapport à un point O d'un point matériel comme le produit vectoriel du vecteur position et du vecteur force :

$$\mathbf{N} = \mathbf{r} \times \mathbf{F} \quad (1a)$$

Le *moment cinétique* \mathbf{L} du point matériel par rapport au point O est défini comme le produit vectoriel du vecteur position et du vecteur quantité de mouvement :

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (2a)$$

C'est la contrepartie angulaire de la quantité de mouvement.

En effectuant le produit vectoriel du vecteur position par l'équation du PFD $\mathbf{F} = \frac{d}{dt}(m\mathbf{v})$, on obtient l'équivalent angulaire du PFD :

$$\mathbf{r} \times \mathbf{F} = \mathbf{N} = \mathbf{r} \times \frac{d}{dt}(m\mathbf{v}) \quad (3a)$$

En employant l'identité vectorielle, l'équation prend la forme :

$$\mathbf{N} = \frac{d}{dt}(\mathbf{r} \times m\mathbf{v}) = \frac{d\mathbf{L}}{dt} \quad (4a)$$

C'est ce qu'on appelle le *théorème du moment cinétique*.

A.2 Le mouvement d'un système de points matériels. Forces externes et internes. Théorème du centre de masse.

Les trois lois de Newton s'étendent sans modification profonde de leur structure aux systèmes de particules. Il faut cependant distinguer entre les *forces externes* agissant sur les points matériels et dues à des causes extérieures au système, et les *forces internes* agissant sur un point matériel i et dues aux autres points matériels du système. Ainsi, la deuxième loi de Newton pour le point i a cette forme :

$$\sum_j \mathbf{F}_{ji} + \mathbf{F}_i^{(e)} = \dot{\mathbf{p}}_i \quad (5a)$$

où $\mathbf{F}_i^{(e)}$ représente une force externe et \mathbf{F}_{ji} est la force interne due à la particule j agissant sur la particule i .

On suppose que $\mathbf{F}_i^{(e)}$, de même que \mathbf{F}_{ji} , obéissent à la loi de l'égalité de l'action et de la réaction. Ce n'est pas le cas pour certaines forces comme, par exemple, les forces électromagnétiques.

En faisant la somme des forces agissant sur toutes les particules, l'équation (5a) devient

$$\frac{d^2}{dt^2} \sum_i m_i \mathbf{r}_i = \sum_i \mathbf{F}_i^{(e)} + \sum_{ij, i \neq j} \mathbf{F}_{ji} \quad (6a)$$

La première somme du second membre est la force externe totale, et la seconde somme est nulle par la loi d'action et réaction. Pour réduire le premier membre, on définit un vecteur \mathbf{R} comme moyenne des vecteurs positions des particules, pondérées proportionnellement à leur masse, à savoir :

$$\mathbf{R} = \frac{\sum m_i \mathbf{r}_i}{\sum m_i} = \frac{\sum m_i \mathbf{r}_i}{M} \quad (7a)$$

Le point défini par le vecteur \mathbf{R} est appelé *centre de masse* du système. Dès lors, l'équation (6a) se réduit à :

$$M \frac{d^2 \mathbf{R}}{dt^2} = \sum_i \mathbf{F}_i^{(e)} \equiv \mathbf{F}^{(e)} \quad (8a)$$

C'est le *théorème du centre de masse*, qui dit que le centre de masse se meut comme si la force externe totale agissait sur la masse entière du système concentrée au centre de masse. Les forces internes n'ont par conséquent aucun effet sur le mouvement du centre de masse.

La quantité de mouvement totale \mathbf{P} d'un système de points matériels est le produit de la masse totale du système et de la vitesse du centre de masse :

$$\mathbf{P} = \sum m_i \frac{d\mathbf{r}_i}{dt} = M \frac{d\mathbf{R}}{dt} \quad (9a)$$

On élargit sans difficulté les notions de moment cinétique et de moment de force ainsi que le théorème du moment cinétique au cas d'un système de particules.

A.3 Les théorèmes de conservation comme conséquences des principes newtoniens

Dans les présentations courantes de la mécanique newtonienne, on ajoute généralement à l'exposé des trois lois du mouvement celui de certains théorèmes de conservation. Ils sont déductibles des principes newtoniens.

Conservation de la quantité de mouvement et du moment cinétique

Ces théorèmes sont des conséquences directes des principes newtoniens. En vertu du PFD, si la force totale \mathbf{F} qui s'applique à un système est nulle, alors la dérivée de la quantité de mouvement \mathbf{p} est nulle et donc \mathbf{p} est constante. C'est le *théorème de conservation de la quantité de mouvement*.

De même, pour des mouvements non rectilignes, le théorème du moment cinétique (4a) permet de déduire le *théorème de conservation du moment cinétique* : si le moment de la force totale \mathbf{N} est nul, alors la dérivée de \mathbf{L} est nulle, et le moment cinétique \mathbf{L} est conservé.

On élargit sans peine ces théorèmes aux systèmes de particules. En effet, l'équation du mouvement pour le centre de masse du système peut être énoncée sous une autre forme comme le théorème de conservation de la quantité de mouvement totale du système : si la force externe totale est nulle, la quantité de mouvement est conservée. On montre ensuite que la dérivée du moment cinétique total est égale au

moment de la force externe par rapport au point donné. Il en résulte le théorème de conservation pour le moment cinétique total, qui s'énonce comme suit : le moment cinétique total est constant dans le temps si le moment de la force totale externe est nul.

Ces théorèmes de conservation, comme on le voit, ne sont que des formulations différentes du PFD. Ils sont utiles cependant lors de l'étude de certains phénomènes mécaniques pour lesquels l'application stricte du PFD ne serait pas d'un grand recours, comme par exemple l'étude des collisions.

Le théorème du moment cinétique est de plus une loi très générale et qui touche à la structure même des lois physiques ; on montre en effet que la conservation du moment cinétique pour un système isolé est due à l'invariance des lois physiques vis-à-vis d'une rotation géométrique des systèmes de coordonnées. Aussi retrouve-t-on la notion de moment cinétique dans d'autres domaines de la physique et dans d'autres formalismes.

Conservation de l'énergie

Contrairement aux théorèmes de conservation de la quantité de mouvement et du moment cinétique, les théorèmes concernant la conservation de l'énergie nécessitent l'introduction de notions qui ne figurent pas dans l'énoncé des principes newtoniens, et qui sont les concepts fondamentaux des formulations analytiques de la mécanique. Dans les exposés de la mécanique newtonienne, ils sont généralement introduits par des définitions *ad hoc*.

Travail d'une force et énergie cinétique

On définit le *travail* W de la force externe \mathbf{F} sur la particule qui se déplace du point 1 au point 2 comme suit

$$W_{12} = \int_1^2 \mathbf{F} \cdot d\mathbf{s} \quad (10a)$$

Pour une masse constante, en intégrant l'équation (10a), on trouve :

$$W_{12} = \frac{m}{2}(\mathbf{v}_2^2 - \mathbf{v}_1^2) \quad (11a)$$

La quantité scalaire $\frac{1}{2}m\mathbf{v}^2$ est appelée *énergie cinétique* de la particule et est notée « T » de telle manière que le travail effectué est égal à la variation de l'énergie cinétique :

$$W_{12} = T_2 - T_1 \quad (12a)$$

Cette dernière équation est appelée « théorème de l'énergie cinétique ».

Remarque : ce que j'appelle théorème ici n'est à la rigueur, ainsi présenté, qu'une définition. En effet, partant de la définition du travail, on a trouvé, en intégrant l'équation (10a), l'expression d'une quantité que l'on a décidé d'appeler « énergie cinétique » ; le théorème de l'énergie cinétique, selon lequel le travail de la force totale qui s'exerce sur un point matériel entre deux instants est égal à la variation de l'énergie cinétique du point entre ces deux mêmes instants n'apporte, en tant que tel, strictement rien. Quand l'énergie devient un concept central, et non plus le nom que l'on donne à une certaine quantité apparue dans un calcul, le théorème peut être considéré comme disant quelque chose de plus que la seule définition de l'énergie.

Energie potentielle et systèmes conservatifs

Le théorème de l'énergie cinétique permet de calculer l'énergie cinétique, donc le carré de la vitesse, si les données concernant le champ de force suffisent à déterminer le travail W . C'est le cas des champs de force dits « conservatifs » ou « dérivant d'un potentiel ».

Un champ de force est dit conservatif quand il est tel que le travail W effectué autour d'une orbite fermée est nul, c'est-à-dire si :

$$\oint \mathbf{F} \cdot d\mathbf{s} = 0 \quad (13a)$$

Les forces de frottement et la résistance de l'air, par exemple, sont des forces non conservatives.

Si l'équation (13a) est vérifiée, le travail doit être indépendant du chemin suivi pour intégrer (10a) entre les points 1 et 2. C'est-à-dire que le travail d'une force conservative le long d'un trajet donné ne dépend que de la position des points extrêmes et non de la trajectoire particulière effectivement suivie. Il s'ensuit qu'il est possible d'exprimer W_{12} comme la variation d'une quantité qui dépende seulement de la position des points terminaux. Cette quantité peut être notée « $-V$ », de telle façon que nous avons la relation $\mathbf{F} \cdot d\mathbf{s} = -dV$, soit $F_s = -\frac{\partial V}{\partial s}$, ce qui revient à dire que \mathbf{F} est le gradient d'un certain scalaire V appelé *potentiel du champ \mathbf{F}* ou *énergie potentielle du point dans le champ de force \mathbf{F}* . Le niveau zéro de cette quantité V est arbitraire.

Pour un système conservatif le travail effectué par les forces est :

$$W_{12} = V_1 - V_2 \quad (14a)$$

En combinant l'équation (14a) et le théorème de l'énergie cinétique (12a) on obtient :

$$T_1 + V_1 = T_2 + V_2, \quad (15a)$$

qui exprime le *théorème de conservation de l'énergie pour une particule* selon lequel si les forces agissant sur une particule sont conservatives, l'*énergie totale de la particule*, $T + V$, est conservée.

Tous ces théorèmes sont élargis sans difficulté au cas d'un système de points matériels.

A.4 La loi de gravitation universelle

La force de gravitation est une force qui s'exerce entre deux corps doués de masse suivant la loi de la gravitation universelle :

$$\mathbf{F} = -G \frac{mm'}{r^2} \mathbf{u}$$

où m et m' sont les masses (gravitationnelles, non distinguées, en mécanique classique, des masses inertes) des deux corps en interaction gravitationnelle, r la distance qui les sépare et G une constante universelle.

On peut tirer de cette loi tout un ensemble d'applications et de lois locales, comme par exemple déterminer la force d'attraction terrestre, approximation obtenue en connaissant la masse de la terre et en considérant comme négligeable la distance à la surface de la terre du corps en chute libre. Ainsi, l'équation d'une particule en chute libre à partir du repos avec une force due à la gravité, s'écrit :

$$g = m \frac{d^2x}{dt^2}$$

A.5 Les différents types de contraintes

Les contraintes peuvent être classées de différentes façons. J'adopte ici le classement proposé par mes deux sources principales, (Goldstein, 1950/2002) et (Lanczos, 1970).

La typologie des contraintes se fait suivant deux distinctions principales :

1. Contraintes holonomes / non holonomes⁵⁹.
 - Les contraintes sont dites *holonomes* quand elles peuvent être exprimées sous forme d'équations reliant les coordonnées des particules et le temps ayant pour forme :

$$f(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, t) = 0 \quad (16a)$$

⁵⁹Cette terminologie est due à Heinrich Hertz (1894).

Les particules d'un corps rigide sont soumises à de telles contraintes :

$$(\mathbf{r}_i - \mathbf{r}_j)^2 - \mathbf{c}_{ij}^2 = 0. \quad (17a)$$

- Les conditions cinématiques ne peuvent pas toujours être exprimées par des équations entre les coordonnées des particules. Ces liaisons ou contraintes sont dites *non holonomes*. Un exemple caractéristique en est le roulement sans glissement d'une balle sur une surface. Les conditions cinématiques d'un tel système ne peuvent être exprimées que par des relations entre les dérivées des coordonnées.
- 2. Contraintes scléronomes / rhéonomes⁶⁰.
 - Les contraintes indépendantes du temps sont appelées *scléronomes*.
 - Les contraintes qui contiennent explicitement le temps sont appelées *rhéonomes*. La relation entre les coordonnées d'un système mécanique soumis à de telles contraintes change alors continuellement avec le temps. Une perle glissante sur un fil en mouvement, un pendule dont le fil change de longueur sont des exemples de liaisons rhéonomes. Je reviendrai sur cette distinction dans la section consacrée à la formulation analytique après avoir exposé les notions d'espace de configuration et de coordonnées généralisées.

A.6 Quelques conséquences du principe de d'Alembert

La conservation de l'énergie

La conservation de l'énergie peut être déduite du principe de d'Alembert dans les cas où la trajectoire virtuelle coïncide avec la trajectoire réelle ; l'équation qui en résulte peut alors être intégrée par rapport au temps. La validité de cette procédure dépend de la constance des masses des particules et du fait que la fonction U et les conditions cinématiques soient scléronomes, c'est-à-dire indépendantes du temps.

Le principe de d'Alembert implique généralement des forces de nature polygénique. Cependant, dans un cas particulier, il devient monogénique et intégrable, et permet alors de déduire la loi de conservation de l'énergie. Si les forces appliquées sont monogéniques, leur travail égale l'opposé de la variation de l'énergie potentielle, et le principe exprimé par l'équation du principe de d'Alembert $\sum (\mathbf{F}_i - m_i \mathbf{a}_i) \cdot \delta \mathbf{r}_i = 0$ devient :

$$\delta V + \sum_i m_i \ddot{\mathbf{r}}_i \cdot \delta \mathbf{r}_i = 0 \quad (18a)$$

⁶⁰Cette terminologie est due à Boltzmann (1897).

Le second terme, l'opposé du travail des forces d'inertie, ne peut généralement pas être exprimé comme la variation d'une quantité. Mais, si l'on fait coïncider les déplacements virtuels $\delta \mathbf{R}_k$ avec les déplacements réels qui ont lieu pendant l'intervalle de temps dt , alors les changements virtuels de l'énergie potentielle coïncideront avec les changements réels, et le second terme devient aussi une dérivée parfaite :

$$\sum m_k \ddot{\mathbf{R}}_k \cdot d\mathbf{R}_k = \sum m_k \ddot{\mathbf{R}}_k \cdot \dot{\mathbf{R}}_k dt = \frac{d}{dt} \left(\frac{1}{2} \sum m_k \dot{\mathbf{R}}_k^2 \right) dt = dT \quad (19a)$$

car $\frac{1}{2} \sum m_k \dot{\mathbf{R}}_k^2 = \frac{1}{2} \sum m_k \mathbf{v}_k^2 = T$, T étant l'énergie cinétique du système mécanique. L'équation (18a) prend alors la forme

$$dV + dT = d(V + T) = 0, \quad (20a)$$

ce qui peut être intégré et donner

$$T + V = \text{constante} = E \quad (21a)$$

La somme des énergies potentielle et cinétique reste la même pendant le mouvement. C'est la loi de conservation de l'énergie.

Dans un espace de configuration à n dimensions correspondant aux n degrés de liberté d'un système, et dont les coordonnées généralisées sont par conséquent indépendantes, on peut toujours faire coïncider les changements virtuels de coordonnées avec les changements réels ; mais ce n'est pas le cas en coordonnées rectangulaires puisque, précisément, elles ne sont pas indépendantes. Dans le cas de contraintes scléronomes, indépendantes du temps, on peut passer de l'identification des δq_i aux dq_i à celle des $\delta \mathbf{R}_i$ aux $d\mathbf{R}_i$; dans le cas de contraintes rhéonomes, ce n'est pas possible.

Une remarque similaire vaut pour l'énergie potentielle V . Si V est une fonction des q_i uniquement, alors l'identification des δq_i aux dq_i conduit à celle de δV à dV . Mais cela ne vaut pas si V dépend explicitement du temps. Puisque V est définie comme $-U$, cela signifie que la fonction U ne doit pas dépendre explicitement du temps pour que la loi de conservation de l'énergie soit respectée.

On voit donc que la loi de conservation de l'énergie est restreinte aux systèmes scléronomes aussi bien quant à leur fonction U que quant à leurs contraintes. On a supposé aussi tout au long de ce raisonnement que les masses sont constantes. Pour plus de détails, voir (Lanczos, 1970, pp. 94-96).

Forces apparentes dans un système de référence en mouvement

Le principe de d'Alembert permet l'usage de systèmes de référence en mouvement. Il affirme que la force d'inertie est une force de plus qui agit exactement comme les

autres. Cependant, pour la définition de la force d'inertie on doit faire usage d'un « système de référence absolu ». Si le système de référence dans lequel les accélérations sont mesurées est en mouvement relativement au système absolu, la force d'inertie mesurée dans le système en mouvement n'est plus la vraie force d'inertie mais doit être corrigée de termes additionnels. Selon le principe de d'Alembert, les termes venant corriger la force d'inertie selon le mouvement du système de référence peuvent indifféremment être interprétés comme appartenant aux forces appliquées produites par un champ externe de forces. Cela donne naissance au phénomène des « forces apparentes » qui se produit dans des systèmes de référence en mouvement. Si un système de référence est en mouvement accéléré, ce mouvement produit le même effet que si des forces externes étaient présentes qui devaient être ajoutées aux forces appliquées. Ces forces sont appelées « forces de d'Alembert » ou « forces apparentes » parce qu'elles sont générées par le mouvement du système de référence mais ne sont pas présentes dans un système absolu. Si le système de référence est dans un état de translation uniformément accélérée, la force apparente générée est celle d'un champ de gravitation uniforme. L'hypothèse d'équivalence introduite par Einstein pour résoudre l'énigme de la gravité avant d'établir le principe général de la relativité est, pour les phénomènes purement mécaniques, une conséquence directe du principe de d'Alembert. Pour plus de détails, voir (Lanczos, 1970, pp. 100-103).

Le principe de moindre contrainte de Gauss

Le principe de moindre contrainte de Gauss est une reformulation du principe de d'Alembert sous la forme d'un principe de minimum. Une quantité scalaire, appelée « contrainte » par Gauss, est minimisée, en considérant les accélérations comme les variables du problème de minimum. Contrairement aux autres principes de minimum, il ne requiert pas le calcul variationnel, parce que c'est une fonction ordinaire, et non une intégrale, qui doit être minimisée. Le gros inconvénient de cette procédure est qu'elle nécessite que l'on connaisse les accélérations. Voir (Lanczos, 1970, pp. 109-110) pour l'interprétation géométrique que Hertz (1894) donne du principe de Gauss, le mettant en relation étroite avec le principe de Jacobi. Pour plus de détails sur le principe de moindre contrainte, voir (Lanczos, 1970, pp. 106-110).

A.7 Dédution des équations de Lagrange à partir du principe de d'Alembert

Les équations lagrangiennes du mouvement sont des conséquences du principe de d'Alembert, exprimé au moyen de coordonnées généralisées. Prenons le principe de

d'Alembert :

$$\sum_i (\mathbf{F}_i^{(a)} - \dot{\mathbf{p}}_i) \cdot \delta \mathbf{r}_i = 0 \quad (22a)$$

Tel quel, il ne permet pas encore de formuler les équations du mouvement ; il faut auparavant le transformer en une expression portant sur les déplacements virtuels des coordonnées généralisées qui sont indépendantes les unes des autres (pour les liaisons holonomes) de telle manière que les coefficients des δq_i puissent être séparément égaux à zéro.

Le passage des coordonnées rectangulaires aux coordonnées généralisées est assuré par les équations de transformation présentées plus haut et achevé au moyen des règles usuelles des dérivées partielles. Ainsi, $\mathbf{v}_i \equiv \dot{\mathbf{r}}_i$ s'écrit

$$\mathbf{v}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \mathbf{r}_i}{\partial t} \quad (23a)$$

Ainsi, le déplacement virtuel $\delta \mathbf{r}_i$ peut être associé aux déplacements virtuels δq_i par

$$\delta \mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \quad (24a)$$

Aucune variation de temps n'est mise en jeu, puisqu'un déplacement virtuel n'implique que des déplacements de coordonnées. En termes de coordonnées généralisées, le travail virtuel des \mathbf{F}_i devient

$$\begin{aligned} \sum_i \mathbf{F}_i \cdot \delta \mathbf{r}_i &= \sum_{i,j} \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \\ &= \sum_i \mathbf{Q}_i \delta q_i \end{aligned} \quad (25a)$$

où les \mathbf{Q}_i sont les composantes de la *force généralisée*, définie comme

$$\mathbf{Q}_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \quad (26a)$$

L'autre terme de l'équation (22a) peut être mis sous la forme :

$$\sum_i \dot{\mathbf{p}} \cdot \delta \mathbf{r}_i = \sum_i m_i \ddot{\mathbf{r}}_i \cdot \delta \mathbf{r}_i \quad (27a)$$

qui devient, en exprimant $\delta \mathbf{r}_i$ par (24a) :

$$\sum_{i,j} m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \quad (28a)$$

Considérons maintenant l'égalité

$$\sum_{i,j} m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_i} = \sum_i \left\{ \frac{d}{dt} (m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_i}) - m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_i} \right) \right\} \quad (29a)$$

Dans le dernier terme de (29a), on peut interchanger la différentiation par rapport à t et à q_j , car, par analogie avec (23a),

$$\frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial q_i} = \sum_k \frac{\partial^2 \mathbf{r}_i}{\partial q_i \partial q_k} \dot{q}_k + \frac{\partial^2 \mathbf{r}_i}{\partial q_i \partial t}, \quad (30a)$$

qui, par (23a) est précisément $\frac{\partial \mathbf{v}_i}{\partial q_j}$. En outre, par (23a),

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_i} = \frac{\partial \mathbf{r}_i}{\partial q_i} \quad (31a)$$

En portant ces égalités dans (29a), on obtient

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_i} = \sum_i \left\{ \frac{d}{dt} (m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_i}) - m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial q_i} \right\} \quad (32a)$$

et le terme en question de l'équation (22a) devient

$$\sum_j \left\{ \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}_j} \left(\sum_i \frac{1}{2} m_i v_i^2 \right) \right) - \frac{\partial}{\partial q_j} \left(\sum_i \frac{1}{2} m_i v_i^2 \right) \right\} \delta q_j \quad (33a)$$

En identifiant $\sum_i \frac{1}{2} m_i v_i^2$ à l'énergie cinétique T du système, le principe de d'Alembert devient :

$$\sum [\left\{ \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} \right\} - \mathbf{Q}_i] \delta q_i = 0 \quad (34a)$$

Or, en vertu du caractère holonome des contraintes, les q_i sont indépendants. Tout déplacement virtuel δq_j est indépendant de δq_k et, par conséquent, la seule façon de garder la validité de (34a) est d'égaliser à zéro séparément les coefficients

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} = \mathbf{Q}_i \quad (35a)$$

ces équations étant au nombre de n .

Les équations (35a) sont souvent appelées les *équations de Lagrange* mais cette désignation est plutôt réservée à la forme de ces équations quand le système est conservatif, c'est-à-dire quand les forces dérivent d'une fonction potentielle scalaire, V , qui est l'énergie potentielle du système :

$$\mathbf{F}_i = -\nabla_i V \quad (36a)$$

Alors, les forces généralisées peuvent s'écrire

$$\mathbf{Q}_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = - \sum_i \nabla_i V \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \quad (37a)$$

ce qui est l'expression de la dérivée partielle d'une fonction $-V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ par rapport à q_j

$$\mathbf{Q}_j = - \frac{\partial V}{\partial q_j} \quad (38a)$$

Ainsi, l'équation (34a) peut s'écrire

$$\frac{d}{dt} \left(\frac{\partial (T - V)}{\partial \dot{q}_j} \right) - \frac{\partial (T - V)}{\partial q_j} = 0 \quad (39a)$$

soit, en définissant une nouvelle fonction, le lagrangien L , comme

$$L = T - V \quad (40a)$$

les équations (35a) prennent la forme

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = 0 \quad (41a)$$

Ce sont les *équations de Lagrange*.

Si le système n'est pas conservatif, les équations de Lagrange peuvent être mises sous la forme 41a, pourvu que les forces généralisées dérivent d'une fonction $U(q_j, \dot{q}_j)$ définie par l'équation

$$\mathbf{Q}_j = - \frac{\partial U}{\partial q_j} + \frac{d}{dt} \left(\frac{\partial U}{\partial \dot{q}_j} \right) \quad (42a)$$

Dans un tel cas, l'équation (41a) résulte des équations (35a) avec le Lagrangien donné par $L = T - U$ où U peut être appelé « potentiel généralisé ».

Il est bon de noter que si quelques unes seulement des forces agissant sur le système sont dérivables d'un potentiel, les équations de Lagrange peuvent toujours être écrites sous la forme

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = Q_j \quad (43a)$$

où L contient le potentiel des forces conservatives et Q_j représente les forces ne dérivant pas d'un potentiel.

A.8 Énergie cinétique, énergie potentielle et travail en mécanique analytique

Énergie cinétique

Dans le cadre de la présentation analytique de la mécanique, la quantité fondamentale qui caractérise l'inertie de la masse n'est pas la quantité de mouvement $m\mathbf{v}$

mais l'énergie cinétique. L'énergie cinétique est une quantité scalaire définie, pour une particule unique, par

$$T = \frac{1}{2} m \mathbf{v}^2 \quad (44a)$$

et, pour un système de particules, par

$$T = \frac{1}{2} \sum_{i=1}^n m_i \mathbf{v}_i^2 \quad (45a)$$

\mathbf{v} est la vitesse de la particule, définie par

$$\mathbf{v}^2 = \left(\frac{ds}{dt} \right)^2 = \frac{dx^2 + dy^2 + dz^2}{dt^2} \quad (46a)$$

Or ds est l'élément de courbe⁶¹ d'un espace riemanien⁶². La quantité T contient donc l'expression d'un tenseur métrique⁶³ dans l'espace de configuration à $3n$ dimensions :

$$ds^2 = 2T dt^2 = \sum_i m_i \mathbf{v}_i^2 dt^2 = \sum_i m_i (dx_i^2 + dy_i^2 + dz_i^2) \quad (47a)$$

L'énergie cinétique totale du système peut alors être écrite sous la forme :

$$T = \frac{1}{2} m \left(\frac{ds}{dt} \right)^2 \quad (48a)$$

avec $m = 1$. Cela signifie que l'énergie cinétique du système entier peut être remplacée par l'énergie cinétique d'un point unique de masse 1. Il s'agit d'un point de l'espace de configuration à $3n$ dimensions qui représente la position du système mécanique. En choisissant un espace de configuration doté du nombre de dimensions adéquat, on peut donc représenter le mouvement de n'importe quel système par celui d'un point unique.

⁶¹ Il s'agit d'une quantité différentielle qui représente la distance entre deux points de l'espace, exprimée au moyen de leurs coordonnées et de leur différentiels.

⁶² Pour plus de détails concernant la géométrie riemannienne et son lien à l'invariance des équations de la mécanique analytique relativement à des changements arbitraires de coordonnées, voir (Lanczos, 1970, pp. 17-21).

⁶³ Voir également (Lanczos, 1970, pp. 19-22). En bref, la notion de tenseur métrique est une généralisation du théorème de Pythagore. Un tenseur est défini par les composantes d'une forme différentielle invariante. Un vecteur, défini par une forme différentielle invariante de premier ordre $dw = F_1 dx_1 + F_2 dx_2 + \dots + F_n dx_n$ est un tenseur de premier ordre. Les quantités F_1, F_2, \dots, F_n sont les composantes du vecteur ; elles changent avec le système de référence utilisé et sont donc des quantités « covariantes ». Un tenseur métrique est un tenseur d'ordre 2. Il permet le développement d'une géométrie complète d'un ordre quelconque de dimensions.

Énergie potentielle et travail, conservation de l'énergie.

Supposons que chaque particule de masse m_i et de coordonnées rectangulaires x_i, y_i, z_i est mue par un ensemble de forces \mathbf{G}_i dont les composantes sont X_i, Y_i, Z_i . Ces forces viennent soit d'un champ externe soit de l'interaction des particules, mais elles n'incluent pas les forces qui maintiennent les contraintes. Changeons les coordonnées de chacune des particules d'une quantité infinitésimale dx_i, dy_i, dz_i . Le travail total effectué par les forces appliquées pour un changement infinitésimal de la position du système peut être décrit par une forme différentielle en coordonnées cartésiennes :

$$dw = \sum_{i=1}^n (X_i dx_i + Y_i dy_i + Z_i dz_i). \quad (49a)$$

ou, si l'on transforme les coordonnées rectangulaires en coordonnées généralisées, par une forme différentielle linéaire des variables q_i :

$$dw = F_1 dq_1 + F_2 dq_2 + \dots + F_n dq_n \quad (50a)$$

Les coefficients F_1, F_2, \dots, F_n prennent la place des composantes originales des forces qui agissent sur chaque particule et sont de première importance : ils sont les composantes d'un vecteur de l'espace de configuration à n dimensions que l'on appelle la « force généralisée » et qui suffit à représenter l'action dynamique de toutes les forces appliquées au système (de même que le mouvement du système peut être représenté comme le mouvement d'une particule unique dans cet espace à n dimensions).

Dans les problèmes de mécanique analytique, dw est le plus souvent le différentiel d'une certaine fonction⁶⁴, que l'on désigne⁶⁵ par la lettre U . Dans un tel cas,

$$dw = dU \quad (51a)$$

Dans les cas les plus simples où U est indépendante du temps t et des vitesses généralisées \dot{q}_i , on a

$$U = U(q_1, q_2, \dots, q_n) \quad (52a)$$

Ainsi,

$$\sum F_i dq_i = \sum \frac{\partial U}{\partial q_i} dq_i \quad (53a)$$

et donc

$$F_i = \frac{\partial U}{\partial q_i} \quad (54a)$$

⁶⁴Dans le cas contraire, on devrait en principe, suivant la convention orthographique adoptée par Lanczos (1970, p. 18, n. 1), écrire « \overline{dw} », la barre indiquant qu'il s'agit d'une forme différentielle non intégrable.

⁶⁵Cette fonction porte en anglais le nom de « *work function* », c'est-à-dire « fonction de travail », mais on lui donne rarement ce nom en français.

En posant $V = -U$, on peut réécrire l'équation (54a) ainsi :

$$F_i = -\frac{\partial V}{\partial q_i} \quad (55a)$$

La force décrite dans l'équation (55a) est une force conservative et, en tant que telle, satisfait la loi de conservation de l'énergie. En outre, la force F_i peut être dérivée de l'unique fonction scalaire U : on dit qu'elle *dérive d'un potentiel*.

Signalons qu'il existe des forces dérivables d'une fonction U *dépendante du temps* (ce qui n'est pas le cas de la force décrite ci-dessus, où U est indépendante du temps t et des vitesses généralisées \dot{q}_i) ; pour de telles forces, l'équation (51a) vaut toujours, si pour former le différentiel dU le temps t est considéré comme une constante. Les équations (53a) et (54a) restent aussi valables, mais la conservation de l'énergie est perdue. La force généralisée dérive alors d'une quantité scalaire sans être conservative. Dans la terminologie adoptée par (Lanczos, 1970, p. 30), les forces dérivant d'un potentiel, qu'elles soient conservatives ou non, sont appelées *monogéniques* ; les forces qui ne dérivent pas d'un potentiel (comme par exemple les forces de frottement), sont dites *polygéniques*. On s'intéressera essentiellement ici aux forces monogéniques.⁶⁶

Ainsi, parmi les forces monogéniques – qui dérivent d'un potentiel –, celles qui sont indépendantes du temps sont dites « conservatives » parce qu'elles satisfont la loi de conservation de l'énergie, selon laquelle une certaine quantité, l'énergie totale du système, reste constante pendant le mouvement. L'énergie totale du système est définie comme la somme de l'énergie cinétique T et de l'énergie potentielle V , avec

$$V = \sum \frac{\partial U}{\partial \dot{q}_i} \dot{q}_i - U \quad (56a)$$

Dans les cas les plus simples où la fonction U est non seulement indépendante du temps t mais également des vitesses généralisées \dot{q}_i , on a plus simplement

$$V = -U \quad (57a)$$

On peut alors considérer l'énergie cinétique T et l'énergie potentielle V comme les deux scalaires fondamentaux de la mécanique, en se passant de U . Dans les cas plus généraux, cependant, la quantité fondamentale n'est pas l'énergie potentielle V mais la fonction U ; quand T et U sont toutes deux indépendantes du temps, les équations du mouvement conduisent au théorème de conservation de l'énergie.

⁶⁶Il existe un autre cas où la conservation de l'énergie est perdue : celui des systèmes dits « rhéonomes » (voir ci-dessus, annexe A.5), où les contraintes sont dépendantes du temps – et où le temps t apparaît par conséquent dans les équations de transformation (12). Dans ce cas également, l'énergie cinétique et la fonction U contiennent le temps t explicitement. Voir (Lanczos, 1970, pp. 31-34).

A.9 Dédution du principe de Hamilton à partir du principe de d'Alembert

Voici comment s'opère la déduction de la version hamiltonienne du principe de moindre action à partir du principe de d'Alembert⁶⁷ : on multiplie $\overline{\delta w^e}$ par dt et on intègre entre les limites $t = t_1$ et $t = t_2$:

$$\int_{t_1}^{t_2} \overline{\delta w^e} dt \equiv \int_{t_1}^{t_2} \sum [\mathbf{F}_i - \frac{d}{dt}(m_i \mathbf{v}_i)] \cdot \delta \mathbf{R}_i dt \quad (58a)$$

On sépare la branche de droite en deux parties ; la première peut s'écrire :

$$\int_{t_1}^{t_2} \sum \mathbf{F}_i \cdot \delta \mathbf{R}_i dt = - \int_{t_1}^{t_2} \delta V dt = - \delta \int_{t_1}^{t_2} V dt \quad (59a)$$

On a supposé pour cela que la fonction U était indépendante des vitesses et on l'a remplacée par $-V$. La partie de la branche de droite peut être intégrée par parties :

$$- \int_{t_1}^{t_2} \frac{d}{dt}(m_i \mathbf{v}_i) \cdot \delta \mathbf{R}_i dt = - \int_{t_1}^{t_2} \frac{d}{dt}(m_i \mathbf{v}_i \cdot \delta \mathbf{R}_i) dt + \int_{t_1}^{t_2} m_i \mathbf{v}_i \cdot \frac{d}{dt}(\delta \mathbf{R}_i) dt \quad (60a)$$

La première partie de cette somme est intégrable :

$$- [m_i \mathbf{v}_i \cdot \delta \mathbf{R}_i]_{t_1}^{t_2} \quad (61a)$$

et la seconde partie, en vertu de la nature interchangeable de la variation et de la différentiation⁶⁸, peut s'écrire

$$\int_{t_1}^{t_2} m_i \mathbf{v}_i \cdot \frac{d}{dt} \delta \mathbf{R}_i dt = \int_{t_1}^{t_2} m_i \mathbf{v}_i \cdot \delta \mathbf{v}_i dt = \frac{1}{2} \int_{t_1}^{t_2} m_i \delta(\mathbf{v}_i \cdot \mathbf{v}_i) dt = \frac{1}{2} \delta \int_{t_1}^{t_2} m_i v_i^2 dt \quad (62a)$$

Si l'on fait la somme pour toutes les particules, on obtient :

$$\int_{t_1}^{t_2} \overline{\delta w^e} dt = \delta \int_{t_1}^{t_2} \frac{1}{2} \sum m_i v_i^2 dt - \delta \int_{t_1}^{t_2} V dt - [\sum m_i \mathbf{v}_i \cdot \delta \mathbf{R}_i]_{t_1}^{t_2} \quad (63a)$$

On introduit l'énergie cinétique T du système mécanique et on définit le lagrangien L :

$$L = T - V \quad (64a)$$

Avec ces définitions, on peut écrire l'équation (63a) sous la forme :

$$\int_{t_1}^{t_2} \overline{\delta w^e} dt = \delta \int_{t_1}^{t_2} L dt - [\sum m_i \mathbf{v}_i \cdot \delta \mathbf{R}_i]_{t_1}^{t_2} \quad (65a)$$

⁶⁷Cette déduction est tirée de Lanczos (1970).

⁶⁸Pour la démonstration de cette propriété, voir Lanczos (1970).

Il faut maintenant que les δR_i s'annulent aux limites t_1 et t_2 . Cela signifie que la position du système mécanique est supposée donnée pour $t = t_1$ et $t = t_2$, et qu'aucune variation n'a lieu à ces limites. On varie entre des limites données. Dans ce cas, le terme limite à droite s'annule et l'intégrale par rapport au temps du travail virtuel des forces appliquées devient la variation d'une intégrale définie :

$$\int_{t_1}^{t_2} \overline{\delta w^e} dt = \delta \int_{t_1}^{t_2} L dt = \delta A, \quad (66a)$$

avec

$$A = \int_{t_1}^{t_2} L dt \quad (67a)$$

Puisque le principe de d'Alembert exige l'annulation de $\delta \bar{w}^e$ à tout instant, la partie gauche de (66a) s'annule aussi. Ainsi, le principe de d'Alembert peut être reformulé comme suit :

$$\delta A = 0 \quad (68a)$$

C'est le principe de Hamilton⁶⁹, qui établit que le mouvement d'un système mécanique est tel que l'intégrale A devient stationnaire pour toute variation arbitraire possible de la configuration du système, pourvu que les configurations initiale et finale soient données.

A.10 Dédution des équations de Lagrange à partir du principe de Hamilton.

Considérons deux trajectoires possibles entre q_1 et q_2 .⁷⁰ L'une, qu'on notera simplement $q_i(t)$, est la trajectoire effectivement suivie. L'autre, qu'on appellera « trajectoire variée », infiniment proche, correspond à chaque instant aux positions $q_i(t) + \delta q_i(t)$, où $\delta q_i(t)$ est un accroissement infinitésimal de la position. Ces deux trajectoires doivent obéir aux mêmes conditions initiales et finales. On a donc $\delta q(1) = \delta q(2) = 0$. On suppose que les q_i et les δq_i sont deux fois différentiables. Le fait que les q_i donnent la trajectoire effectivement suivie a pour conséquence que l'action A est extrémale sur cette trajectoire et ne varie donc pas au premier ordre dans les δq_i quand on passe à la trajectoire variée. La variation de l'action s'écrit :

$$\delta A = \int_{t_1}^{t_2} (L(q_i + \delta q_i, \dot{q}_i + \delta \dot{q}_i, t) - L(q_i, \dot{q}_i, t)) dt \quad (69a)$$

⁶⁹Le nom « principe de Hamilton », donné par Jacobi, a été adopté au cours du XX^e siècle.

⁷⁰La déduction qui suit est empruntée au cours de Jean-Michel Raimond (2000).

En développant L au premier ordre dans les δq_i , on a :

$$\delta A = \sum_i \int_{t_1}^{t_2} \frac{\partial L}{\partial q_i} \delta q_i dt + \sum_i \int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \frac{d\delta q_i}{dt} dt \quad (70a)$$

Pour fixer la trajectoire, on recherche une condition sur les δq_i . Pour éliminer leurs dérivées temporelles dans l'expression précédente on intègre les termes de la seconde somme par parties. On obtient :

$$\delta A = \sum_i \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] \delta q_i dt + \left[\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right]_{t_1}^{t_2} \quad (71a)$$

Les accroissements infinitésimaux δq_i s'annulant aux extrémités de la trajectoire, le terme tout intégré est identiquement nul. La somme, elle, ne peut s'annuler pour des δq_i arbitraires et indépendants que si les n équations différentielles

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \quad (1 \leq i \leq n) \quad (72a)$$

Ce sont les équations lagrangiennes du mouvement.

A.11 Théorèmes de conservation et propriétés de symétrie

Les n équations lagrangiennes du mouvement pour un système à n degrés de liberté sont des équations différentielles du second ordre par rapport au temps. Dans de très nombreux cas, ces équations ne sont pas complètement intégrables. Il est cependant possible d'accomplir une intégration partielle permettant d'obtenir des équations différentielles du premier ordre de la forme

$$f(q_1, q_2, \dots, \dot{q}_1, \dot{q}_2, \dots, t) = \text{constante} \quad (73a)$$

appelées « intégrales premières des équations du mouvement ». Ces équations fournissent des renseignements sur la dynamique du système étudié, même si sa trajectoire n'est pas totalement connue. Elles permettent notamment d'établir un lien entre les propriétés du système et les lois de conservation présentées ci-dessus (annexe A.3).

Les intégrales premières du mouvement correspondent en effet aux quantités conservées au cours du mouvement. Il existe en général une quantité conservée associée à chaque invariance du système sous une transformation donnée (à chaque propriété de symétrie). Dans le cas d'un système où le lagrangien est invariant par translation dans le temps (par exemple un système isolé, où le lagrangien ne peut dépendre explicitement du temps), l'énergie est une quantité conservée ; de même,

l'invariance par translation dans l'espace implique la conservation de la quantité de mouvement, et l'invariance par rotation celle du moment cinétique.

Le lagrangien est généralement une fonction de toutes les coordonnées de position q_i et de vitesse \dot{q}_i . Il arrive cependant que le lagrangien ne contienne pas une certaine coordonnée q_i (bien qu'il puisse contenir la vitesse correspondante \dot{q}_i) ; cette coordonnée est alors dite « cyclique » ou « ignorable ».

L'équation lagrangienne du mouvement pour une coordonnée cyclique se réduit à

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \quad (74a)$$

C'est la dérivée du lagrangien par rapport aux vitesses généralisées. On pose :

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (75a)$$

p_i étant les composantes du « moment généralisé » (le terme « moment » est synonyme de « quantité de mouvement ») : si l'on s'intéresse à une unique particule libre en utilisant des coordonnées rectangulaires, les trois quantités p_1, p_2, p_3 deviennent identiques aux composantes rectangulaires de la quantité de mouvement $m\mathbf{v}$. Notons cependant que si les q_i ne sont pas des coordonnées rectangulaires, p_i n'a pas forcément la dimension d'une quantité de mouvement.

L'équation de Lagrange pour une coordonnée cyclique se réduit alors à

$$\frac{dp_i}{dt} = 0 \quad (76a)$$

ce qui signifie que p_i est constante au cours du mouvement. Ainsi, on peut énoncer le théorème général selon lequel le moment généralisé conjugué à une coordonnée cyclique est conservé.

Les variables cycliques sont d'une grande importance dans le traitement analytique de la mécanique, et tout particulièrement dans la formulation hamiltonienne⁷¹. Dans le cadre lagrangien, les équations gagnent déjà en simplicité si l'on modifie le lagrangien de telle façon qu'il ne contienne plus que son moment généralisé.

Cette reformulation des équations de Lagrange permet de déduire les théorèmes de conservation des propriétés de symétrie du système. Une coordonnée cyclique correspond à une invariance. Si le système est invariant par translation dans une certaine direction, la quantité de mouvement correspondante est conservée. De même, si une coordonnée de rotation est cyclique (c'est-à-dire si le lagrangien est invariant sous une rotation autour d'un axe), le moment cinétique est conservé.

⁷¹Routh fut le premier à reconnaître l'importance de ces variables en 1877 ; pour le procédé de réduction dite « routhienne », voir (Butterfield, 2004, pp. 78-80).

Enfin, on retrouve la loi de conservation de l'énergie totale pour les systèmes conservatifs holonomes. Le lagrangien ne contient alors pas le temps. On trouve

$$\frac{d}{dt}(L - \sum_j \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}) = 0 \quad (77a)$$

L'expression entre parenthèses est donc une constante que l'on note $-H$. On a alors

$$H = \sum_j \dot{q}_i p_i - L \quad (78a)$$

Pour un système conservatif, on peut montrer que l'intégrale du mouvement $\sum_j \dot{q}_i p_i - L$ est l'énergie totale du système (voir Goldstein, 1950/2002, pp. 61-62) :

$$H = T + V \quad (79a)$$

A.12 Le temps comme variable cyclique ; le principe de Jacobi ; le principe de moindre action

La notion de variable cyclique prend une importance considérable dans les développements ultérieurs de la mécanique. Le procédé de Routh, ainsi que la mécanique sans force proposée par Hertz reposent sur l'utilisation de telles variables. Jacobi en a fait un usage tout à fait remarquable, qui conduit à une nouvelle formulation du principe de moindre action, appelé « principe de Jacobi » ou simplement « principe de moindre action ».

Considérons un système scléronome ou « conservatif » dont le lagrangien ne contient pas le temps explicitement. Imaginons que nous n'utilisons pas le temps comme variable indépendante, mais que les $n + 1$ variables et t sont prises comme fonctions d'un paramètre τ . Le système a donc $n + 1$ degrés de liberté. En désignant les dérivées par rapport à τ par un prime, l'intégrale d'action prend la forme suivante :

$$A = \int_{\tau_1}^{\tau_2} L(q_1, \dots, q_n; \frac{q'_1}{t'}, \dots, \frac{q'_n}{t'}) t' d\tau \quad (80a)$$

Le temps est maintenant une variable cyclique puisque seul t' apparaît et non t . Si l'on utilise le théorème selon lequel le moment conjugué à une variable cyclique est une constante du mouvement, on obtient un théorème qui vaut que le système soit conservatif ou non : le moment associé au temps t est l'opposé de l'énergie totale. Si t est cyclique, c'est-à-dire si notre système est conservatif, on obtient

$$p_t = -\Lambda = \text{constante} = -E \quad (81a)$$

ce qui est une nouvelle dérivation du théorème de l'énergie.

On peut aller plus loin. On sait qu'une variable cyclique peut être éliminée, en réduisant le problème original d'un degré de liberté. Ici, on peut obtenir un nouveau problème variationnel en éliminant t qui détermine le mouvement dans l'espace mais ne dit rien sur la manière dont la trajectoire est décrite relativement au temps.

On modifie d'abord le lagrangien, et notre nouvelle intégrale d'action devient

$$\bar{A} = 2 \int_{\tau_1}^{\tau_2} T t' d\tau \quad (82a)$$

On fait ensuite usage du point C représentant le système mécanique dans l'espace de configuration. On sait que l'énergie cinétique du système peut être écrite comme l'énergie cinétique d'une particule unique de masse 1.

L'élimination de t' du théorème de l'énergie donne finalement

$$\bar{A} = \int_{\tau_1}^{\tau_2} \sqrt{2(E - V)} \frac{ds}{d\tau} d\tau = \int_{\tau_1}^{\tau_2} \sqrt{2(E - V)} d\bar{s} \quad (83a)$$

On a accompli le processus d'élimination et obtenu l'intégrale d'action du système réduit. Le temps t n'apparaît plus dans \bar{A} , qui est de plus indépendant du paramètre τ .

Le principe qui consiste à minimiser l'intégrale (83a) dans le but de trouver la trajectoire du système mécanique est appelé *principe de Jacobi*. Le temps n'apparaît pas dans cette formulation. Il détermine la trajectoire du point C dans l'espace de configuration, et non le mouvement dans le temps.

Le principe de Jacobi est un principe fondamental de la mécanique. Si l'on se restreint au cas d'une particule unique, l'élément de ligne $d\bar{s}$ devient identique à l'élément de ligne d'un espace ordinaire à trois dimensions dans des coordonnées curvilignes arbitraires. Le principe de Jacobi présente donc une remarquable analogie avec le principe du moindre temps de Fermat en optique qui détermine la trajectoire optique en minimisant l'intégrale

$$I = \int_{\tau_1}^{\tau_2} n d\bar{s} \quad (84a)$$

où n est l'indice de réfraction, qui peut changer de point en point, se comportant ainsi de façon similaire à $\sqrt{2(E - V)}$. Cette analogie joue un rôle fondamental dans le développement de la mécanique ondulatoire moderne.

L'analogie ne va pas au delà de la trajectoire décrite par un point en mécanique et un rayon de lumière en optique. La manière dont le mouvement se produit dans le temps est entièrement différente dans les deux cas.

Le principe de Jacobi peut être conçu comme le résultat des étapes suivantes :

1. Remplacer la différenciation par rapport à t par une différenciation par rapport à τ dans l'énergie cinétique

2. Minimiser l'intégrale d'action après avoir éliminé t' par le biais de la relation d'énergie

$$\frac{T'}{t'^2} + V = E. \quad (85a)$$

Ainsi, dans la représentation paramétrique du mouvement le temps est une coordonnée de position supplémentaire qui peut prendre part au processus de variation. Le moment associé au temps est l'opposé de l'énergie totale. Pour des systèmes scléronomes, le temps devient une variable cyclique et le moment correspondant est une constante. Cela conduit au théorème de l'énergie des systèmes conservatifs. L'élimination du temps comme coordonnée ignorable donne un nouveau principe qui détermine uniquement la trajectoire du système mécanique, et pas le mouvement dans le temps. C'est le principe de Jacobi qui est analogue au principe de Fermat en optique. Ce principe peut être appelé de façon générique *principe de moindre action*. Dans ce dernier, l'intégrale de temps du double de l'énergie cinétique est minimisée avec les conditions auxiliaires selon lesquelles les mouvements actuels et variés doivent satisfaire le théorème de l'énergie pendant le mouvement. Si ce principe est traité par la méthode des multiplicateurs de Lagrange, les équations résultantes sont les équations lagrangiennes du mouvement.

En outre, le principe de Jacobi met en évidence la relation intime qui existe entre le mouvement des systèmes conservatifs holonomes et la géométrie des espaces courbes. On peut associer au mouvement d'un système mécanique sous l'action de l'énergie potentielle V le mouvement d'un point le long d'une géodésique d'un espace riemanien donné. Le problème de trouver la solution d'un problème dynamique donné est mathématiquement équivalent au problème de trouver une géodésique. Le point C du système mécanique (quand l'énergie potentielle est nulle) se meut le long de la ligne la plus courte ou géodésique de l'espace de configuration, et à vitesse constante. C'est une généralisation du principe d'inertie.

Notons la relation entre ce principe géodésique et le principe dynamique de la théorie d'Einstein. La loi d'inertie donne le mouvement des planètes sans l'introduction d'une force de gravité. Le principe de Jacobi vaut pour la mécanique relativiste d'une particule. La différence est qu'ici la structure riemannienne de l'espace-temps est une propriété innée du monde et non pas la conséquence de contraintes cinématiques.

Ainsi, le principe de Jacobi relie le mouvement de systèmes conservatifs holonomes avec la géométrie riemannienne. En particulier, si le système se meut par sa propre inertie sans force appliquée, le point C représentant le système décrit une géodésique ou plus court chemin dans l'espace de configuration, qui est un espace riemanien à n dimensions. Le théorème de l'énergie ajoute que le mouvement se produit à vitesse constante. C'est une généralisation naturelle de la loi ordinaire d'inertie qui

établit qu'une particule sous l'action de sa propre inertie a un mouvement rectiligne et uniforme.

Bibliographie

- ACHINSTEIN, P. (1963). Theoretical terms and partial interpretation. *British Journal for the Philosophy of Science*, 14(54):89–105.
- ACHINSTEIN, P. (1965). The problem of theoretical terms. *American Philosophical Quarterly*, 2:193–203.
- ACHINSTEIN, P. (1968). *Concepts of Science : A Philosophical Analysis*. Baltimore : John Hopkins University Press.
- ADAMS, E. W. (1959). The foundations of rigid body mechanics and the derivation of its laws from those of particle mechanics. In HENKIN, L., SUPPES, P. et TARSKI, A., éditeurs : *The Axiomatic Method*. Amsterdam : North-Holland Publishing Co.
- ALLEN, G. (1974). Opposition to the Mendelian-chromosome theory : The physiological and developmental genetics of Richard Goldschmidt. *Journal of the History of Biology*, 7:49–92.
- ALLEN, G. (1978). *Thomas Hunt Morgan*. Princeton : Princeton University Press.
- ANDERSON, J. (1984). Representational types : A tricode proposal. Rapport technique 82.1, Office of Naval Research, Washington, D.C.
- ARNOLD, V. I. (1978). *Mathematical Methods of Classical Mechanics*. New York : Springer-Verlag. Traduction anglaise par K. Vogtmann et A. Weinstein.
- ARNOLD, V. I. (1988). *Geometrical Methods in the Theory of Ordinary Differential Equations*. Berlin : Springer.
- BAIGRIE, B., éditeur (1996). *Picturing Knowledge. Historical and Philosophical Problems Concerning the Use of Art in Science*. Toronto : University of Toronto Press.
- BALZER, W. (1985). On a new definition of theorecity. *Dialectica*, 39:127–145.
- BALZER, W., MOULINES, C. U. et SNEED, J. (1987). *An Architectonic for Science. The Structuralist Program*. Dordrecht : D. Reidel Publishing Company.

- BANKS, E. C. (2003). *Enst Mach's World Elements : A Study in Natural Philosophy* (Western Ontario Series in the Philosophy of Science, volume 68). Berlin : Springer.
- BARBEROUSSE, A. (2008). From one version to the other : intra-theoretical change. In SOLER, L., SANKEY, H. et HOYNINGEN-HUENE, P., éditeurs : *Rethinking Scientific Change and Theory Comparison : Stabilities, Ruptures, Incommensurabilities* (Boston Studies in the Philosophy and History of Science, volume 255), pages 87–101. New York : Springer-Verlag.
- BARBEROUSSE, A. et LUDWIG, P. (2000). Les modèles comme fictions. *Philosophie*, 68:16–43.
- BARROSO FILHO, W. (1994). *La mécanique de Lagrange : principes et méthodes*. Paris : Karthala.
- BARROSO FILHO, W. et COMTE, C. (1988). La formalisation de la dynamique par Lagrange (1736-1813). In RASHED, R., éditeur : *Sciences à l'Epoque de la Revolution*, pages 329–348. Paris : Albert Blanchard.
- BATESON, W. (1906). The progress of genetics since the redicover of Mendel's papers. In LOTSY, J. P., éditeur : *Progressus Rei Botanicae, Association Internationale des Botanites*, pages 368–418. Iena : G. Fischer.
- BATESON, W. (1913). *Problems of Genetics*. New Haven : Yale University Press.
- BATESON, W. (1916). Review of *The Mechanism of Mendelian Heredity*. *Science*, 44.
- BATESON, W. et PUNNETT, R. C. (1908). The heredity of sex. *Science*, 27.
- BATESON, W. et PUNNETT, R. C. (1911a). On gametic series involving reduplication of certain terms. *Journal of Genetics*, 1:293–302.
- BATESON, W. et PUNNETT, R. C. (1911b). On the interrelations of genetic factors. *Proceedings of the Royal Society*, B, 84.
- BATESON, W., PUNNETT, R. C. et SAUNDERS, E. R. (1905). Further experiments on inheritance in sweet peas and stocks : Preliminary account. *Proceedings of the Royal Society*, B, 77.
- BATESON, W. et SAUNDERS, E. R. (1902). The facts of heredity in the light of Mendel's discovery. *Reports to the Evolution Committee of the Royal Society*, 1:125–160.

- BEATTY, J. (1982). What's wrong with the received view of evolutionary biology ? *In* ASQUITH, P. D. et GIERE, R. N., éditeurs : *Proceedings of the 1980 Biennial Meetings of the Philosophy of Science Association*, volume 2. East Lansing, Michigan : Philosophy of Science Association.
- BECKNER, M. (1959). *The biological way of thought*. New York : Columbia University Press.
- BELL, J. (1987). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge : Cambridge University Press.
- BETH, E. (1940). Toward an up-to-date philosophy of the natural sciences. *Methodos*, 1.
- BEURTON, P., FALK, R. et RHEINBERGER, H.-J., éditeurs (2000). *The Concept of Gene in Development and Evolution. Historical and Epistemological Perspectives*. Cambridge : Cambridge University Press.
- BLACHOWICZ, J. (1997). Analog representation beyond mental imagery. *The Journal of Philosophy*, 94(2):55–84.
- BLACK, M. (1962). *Models and Metaphor*. Ithaca, N.Y. : Cornell University Press.
- BLACKMORE, J. T. (1972). *Ernst Mach ; his work, life, and influence*. Berkeley : University of California Press.
- BLACKMORE, J. T. (1995). *Ludwig Boltzmann, His Later Life and Philosophy, 1900-1906. Book Two : The Philosopher*. Dordrecht : Kluwer.
- BLAY, M. (1992). *La naissance de la mécanique analytique*. Paris : PUF.
- BLAY, M. (1993). *Les raisons de l'infini : du monde clos à l'univers mathématique*. Gallimard.
- BLAY, M. (1995). *Les « Principia » de Newton*. Paris : PUF.
- BLAY, M. (2002). *La science du mouvement*. Paris : Belin Sup.
- BLOCK, N. (1981). *Imagery*. Cambridge MA : MIT Press.
- BLOCK, N. (1986). Advertisement for a semantics for psychology. *Midwest Studies in Philosophy*, 10:615–678.
- BLOCK, N. (1987). Functional role and truth conditions. *Proceedings of the Aristotelian Society*, 61:157–181.

- BLOCK, N. (1998). Semantics, conceptual role. In CRAIG, E., éditeur : *Routledge Encyclopedia of Philosophy*.
- BLOOR, D. (1976). *Knowledge and Social Imagery*. Chicago : University of Chicago Press.
- BOGDAN, R., éditeur (1979). *Patrick Suppes*. Dordrecht : D. Reidel Publishing Company.
- BOHR, N. (1961). *Physique atomique et connaissance humaine*. Paris : Gallimard, coll. Folio Essais, 1991. Traduction par E. Bauer et R. Omnès, revue par C. Chevalley. Édition établie par C. Chevalley.
- BOLTZMANN, L. (1897). *Vorlesungen über die Principe der Mechanik*. Johann Ambrosius Barth. Réédité dans (Boltzmann, 1974).
- BOLTZMANN, L. (1905). *Populäre Schriften*. Leipzig : J.A. Barth.
- BOLTZMANN, L. (1974). *Theoretical Physics and Philosophical Problems*. Dordrecht : D. Reidel Publishing Company.
- BONNET, C. et WAGNER, P., éditeurs (2006). *L'âge d'or de l'empirisme logique. Vienne-Berlin-Prague 1929-1936*. Paris : Gallimard.
- BOTTAZINI, U. (1986). *The Higher Calculus : A History of Real and Complex Analysis from Euler to Weierstrass*. New York : Springer-Verlag.
- BOVERESSE, J. (1973). La théorie et l'observation dans la philosophie des sciences du positivisme logique. In CHÂTELET, F., éditeur : *Le XX^e Siècle. Histoire de la philosophie VIII*, pages 76–134. Paris : Hachette.
- BOVERESSE, J. (1991). Hertz, Boltzmann et le problème de la “vérité” des théories. In LICHNÉROWICZ, A. et GADOFFRE, G., éditeurs : *La vérité est-elle scientifique ?*
- BOVERESSE, J. et WAGNER, P., éditeurs (2008). *Mathématiques et expérience (1918-1940). L'empirisme logique à l'épreuve*. Paris : Odile Jacob.
- BRANDON, R. B. (2000). *Articulating Reasons. An Introduction to Inferentialism*. Cambridge MA : Harvard University Press.
- BRANNIGAN, A. (1985). L'obscurcissement de Mendel. In *Les scientifiques et leurs alliés*, pages 53–87. Paris : Pandore.
- BRENNER, A. (2000). *Duhem : Science, réalité et apparence*. Paris : Vrin.

- BRIDGES, C. B. (1913). Non-disjunction of the sex chromosomes of drosophila. *Journal of Experimental Zoology*, 15:587–606.
- BRIDGES, C. B. (1916a). Non-disjunction as a proof of the chromosome theory of heredity. *Genetics*, 1(1):1–52.
- BRIDGES, C. B. (1916b). Non-disjunction as a proof of the chromosome theory of heredity (concluded). *Genetics*, 1(2):107–163.
- BRIDGMAN, P. (1927). *The Logic of Modern Physics*. New York : Macmillan.
- BRIDGMAN, P. (1936). *The Nature of Physical Theory*. Princeton, N.J. : Princeton University Press.
- BUDD, M. (1993). How pictures look. In KNOWLES, D. et SKORUPSKI, J., éditeurs : *Virtue and Taste : Essays on Politics, Ethics, and Aesthetics*. Oxford : Basil Blackwell.
- BURIAN, R. (1985a). Conceptual change, cross-theoretical explanation and the unity of science. *Synthese*, 33:1–28.
- BURIAN, R. (1985b). On conceptual change in biology : The case of the gene. In DEPEW, D. et WEBER, B., éditeurs : *Evolution at a Crossroads : The New Biology and the New Philosophy of Science*, pages 21–42. Cambridge, MA/London : MIT Press.
- BURIAN, R. (2005). Too many kinds of genes ? some problems posed by discontinuities in gene concepts and the continuity of the genetic material. In citetburian2005, pages 166–178.
- BUTTERFIELD, J. (2004). Between laws and models : Some philosophical morals of Lagrangian mechanics. Disponible sur <http://philsci-archive.pitt.edu/archive/00001937/>.
- BUTTERFIELD, J. (2006). On symmetries and conserved quantities in classical mechanics. In DEMOPOULOS, W. et PITOWSKY, I., éditeurs : *Physical Theory and its Interpretation* (Western Ontario Series in Philosophy of Science), pages 43–99. Dordrecht : Springer.
- CALLENDER, C. et COHEN, J. (2006). There is no special problem about scientific representation. *Theoria*, 55:67–85.

- CALLON, M. et LATOUR, B., éditeurs (1991). *La science telle qu'elle se fait. Anthologie de la sociologie des sciences de langue anglaise*. Paris : La Découverte, « Textes à l'appui. Anthropologie des sciences et des techniques ».
- CAMPBELL, N. R. (1920). *Physics : The Elements*. Cambridge University Press. Réédité sous le titre *Foundations of Science*, New York : Dover, 1957.
- CANTOR, G. (1883). Grundlagen einer allgemeinen Mannigfaltigkeitslehre. *Mathematische Annalen*, 21:545–591.
- CAREY, S. (1985). *Conceptual Change in Childhood*. Cambridge MA : MIT Press.
- CAREY, S. (1991). Knowledge acquisition : Enrichment or conceptual change? In CAREY, S. et GELMAN, R., éditeurs : *The Epigenesis of Mind : Essays on Biology and Cognition*. Mahwah NJ : Lawrence Erlbaum Associates.
- CARLSON, E. A. (1967). *The Gene : A Critical History*. Philadelphia : W.B. Saunders, reprinted by the University of Iowa Press.
- CARNAP, R. (1932a). Die physicalische Sprache als Universalsprache der Wissenschaftssenschaft. *Erkenntnis*, 2. Traduction française de D. Chapuis Schmitz, in (Bonnet et Wagner, 2006).
- CARNAP, R. (1932b). Überwindung der Metaphysik durch logische Analyse der Sprache. *Erkenntnis*, 2:219–241. Traduction française par Barbara Cassin et al., « Le dépassement de la métaphysique par l'analyse logique du langage », in (Soulez, 1985), pages 153–179.
- CARNAP, R. (1934a). *Die Aufgabe der Wissenschaftslogik*. Vienne : Gerold, Einheitswissenschaft. Traduction française par S. Colas, D. Chapuis-Schmitz et P. Wagner, « La tâche de la logique de la science », in Laugier S. et Wagner P., éd., *Philosophie des sciences*, vol. 1, Paris : PUF, 2004.
- CARNAP, R. (1934b). *Logische Syntax der Sprache*. Vienne : Julius Springer, Schriften zur wissenschaftlichen Weltauffassung. Traduction anglaise augmentée Amethe Smeaton : *The Logical Syntax of Language*, Londres : Kegan Paul, Trench, Trubner, 1937 ; traduction française Jacques Bouveresse : *La Syntaxe logique du langage*, Paris : Gallimard, à paraître.
- CARNAP, R. (1936/1937). Testability and meaning. *Philosophy of Science*, 3 : 419-471 ; 4 : 1-40.

- CARNAP, R. (1939a). *Foundations of Logic and Mathematics, International Encyclopedia of Unified Science*, volume 1. Chicago : University of Chicago Press.
- CARNAP, R. (1939b). Theories as partially interpreted formal systems. *In* Carnap (1939a). Rééd. in B.A. Brody and R.E. Grandy, *Readings in the Philosophy of Science*, Prentice Hall, 1989.
- CARNAP, R. (1942). *Introduction to Semantics*. Cambridge MA : Harvard University Press.
- CARNAP, R. (1952). Meaning postulates. *Philosophical Studies*. Repris dans (Carnap, 1956a), 3:65–73.
- CARNAP, R. (1955). Meaning and synonymy in natural languages. *Philosophical Studies*, 6(3):33–47.
- CARNAP, R. (1956a). *Meaning and Necessity*. Chicago : University of Chicago Press.
- CARNAP, R. (1956b). The methodological character of theoretical concepts. *In Minnesota Studies in the Philosophy of Science. The Foundations of Science and Concepts of Psychology and Psychoanalysis*, volume 1, pages 38–76. Minneapolis : University of Minnesota Press.
- CARNAP, R. (1966). *Philosophical Foundation of Physics*. New York : Basic Books.
- CARNAP, R., HAHN, H. et NEURATH, O. (1929). Wissenschaftliche Weltauffassung. der Wiener Kreis. *In Veröffentlichung des Vereines Ernst Mach Heft I*. Wien : Artur Wolf Verlag. Traduction française par B. Cassin sous le titre « La conception scientifique du monde. Le Cercle de Vienne » in (Soulez, 1985).
- CARTWRIGHT, N. (1983). *How the laws of physics lie*. Oxford : Clarendon Press.
- CARTWRIGHT, N. (1999). *The Dappled World : A Study of the Boundaries of Science*. Cambridge : Cambridge University Press.
- CASATI, R. et VARZI, A. C. (1999). *Parts and Places*. Cambridge MA : MIT Press.
- CASTLE, W. E. (1913). Simplification of mendelian formulae. *The American Naturalist*, 47:170–182.
- CASTLE, W. E. (1919). Is the arrangement of the genes in the chromosome linear ? *Proceedings of the National Academy of Sciences of the United States of America*, 5(2):25–32.

- CAT, J. (2001). On understanding : Maxwell on the methods of illustration and scientific metaphor. *Studies in History and Philosophy of Modern Physics*, 32(3): 395–441.
- CHAKRAVARTTY, A. (2001). The semantic of model-theoretic view of theories and scientific realism. *Synthese*, 127:325–345.
- CHANGEUX, J.-P., éditeur (2003). *La vérité dans les sciences*. Paris : Odile Jacob.
- CHAPUIS-SCHMITZ, D. (2006). *Le Sens à l'épreuve de l'expérience. Carnap, Schlick et le vérificationnisme*. Thèse de doctorat, Université Paris 1.
- CHEVALLEY, C. (1991). Introduction et glossaire. In Bohr (1961). Paris : Gallimard, collection Folio Essais.
- CHEVALLEY, C. (1998). Introduction. In Heisenberg (1942). Paris : Éditions du Seuil.
- CLARK, A. (1997). *Being There : Putting Brain, Body and World Together Again*. Cambridge MA : MIT Press.
- COCK, A. (1983). William Bateson's rejection and eventual acceptance of chromosome theory. *Annals of Science*, 40:19–60.
- COHEN, I. B. (1971). *Introduction to Newton's "Principia"*. Cambridge MA : Harvard University Press.
- COHEN, I. B. (1980). *The Newtonian Revolution*. Cambridge : Cambridge University Press.
- COHEN, J. (2004). Information and content. In FLORIDI, L., éditeur : *The Blackwell Guide to the Philosophy of Computing and Information*, pages 215–227.
- COHEN, R. et SEEGER, R., éditeurs (1970). *Ernst Mach – Physicist and Philosopher*. Dordrecht : D. Reidel Publishing Company.
- COLEMAN, W. (1970). Bateson and chromosomes : Conservative thoughts in science. *Centaureus*, 15:228–314.
- CORRENS, C. (1900). G. Mendels Regel über das Verhalten der Nachkommenschaft der Rassenbastarde. *Berichte der Deutschen Botanischen Gesellschaft*, 18:158–168.
- CREIGHTON, H. B. et MCCLINTOCK, B. (1931). A correlation of cytological and genetical crossing-over in *Zea Mays*. *Proceedings of the National Academy of Sciences of the United States of America*, 17:492–497.

- CRICK, F. (1970). Central dogma of molecular biology. *Nature*, 227:561–563.
- CROMBIE, A. C. (1994). *Styles of Scientific Thinking in the European Tradition : The history of argument and explanation especially in the mathematical and biomedical sciences and arts*. London : Duckworth.
- D'AGOSTINO, S. (1990). Boltzmann and Hertz on the Bild-conception of physical theory. *History of Science*, 28:380–398.
- DAHAN-DALMEDICO, A. (1988). Méthodes et « styles » de mathématisation : la science de l'élasticité. In RASHED, R., éditeur : *Sciences à l'Epoque de la Revolution*, pages 349–442. Paris : Albert Blanchard.
- DAHAN-DALMEDICO, A. (1990). Le formalisme variationnel dans les travaux de Lagrange. In LICHNEROWICZ, M., éditeur : *La « Mécanique Analytique » de Lagrange et son héritage*. Actes du Colloque de la Fondation Hugot (septembre 1988), pages 81–106. Turin : Presses de l'Accademia delle scienze di Torino.
- D'ALEMBERT, J. (1743). *Traité de dynamique*. Paris : Jacques Gabay, 1990, fac-similé de l'édition de 1758, revue et augmentée par l'auteur, chez David, Paris.
- DARDEN, L. (1977). William Bateson and the promise of mendelism. *Journal of the History of Biology*, 10:87–106.
- DARDEN, L. (1991). *Theory Change in Science : Strategies from Mendelian Genetics*. Oxford : Oxford University Press.
- DARDEN, L. et MAULL, N. (1977). Interfield theories. *Philosophy of Science*, 44(1):43–64.
- DARRIGOL, O. (2001). God, waterwheels, and molecules : Saint-Venant's anticipation of energy conservation. *Historical Studies in the Physical and Biological Sciences*, 31:285–353.
- DARRIGOL, O. (2007). On the necessary truth of the laws of classical mechanics. *Studies in the History and Philosophy of Modern Physics*, 38:757–800.
- DARRIGOL, O. (2008). The modular structure of physical theories. *Synthese*, 162:195–223.
- DARWIN, C. (1875). *The Variation of Animals and Plants under Domestication*. London : Murray. (2^e édition. 1^{re} édition : 1868).

- DAVIDSON, D. (1974). On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association*, 47:5–20.
- de COURTENAY, N. (1999). *Science et philosophie de Ludwig Boltzmann. La liberté des images par les signes*. Thèse de doctorat, Université Paris 4.
- DE GANDT, F. (1995). *Force and Geometry in Newton's "Principia"*. Princeton : Princeton University Press. Trad. C. Wilson.
- de MAUPERTUIS, P.-L. M. (1746). Les lois du mouvement et du repos déduites d'un principe métaphysique. *Histoire de l'Académie Royale des Sciences de Paris pour l'année 1746*. 1748.
- de REGT, H. W. (1999). Ludwig Boltzmann's *Bildtheorie* and scientific understanding. *Synthese*, 119:113–134.
- de REGT, H. W. et DIEKS, D. (2005). A contextual approach to scientific understanding. *Synthese*, 144(1):137–170.
- de VRIES, H. (1889). *Intracellulare Pangenesis*. Iena : Gustav Fischer. Trad. Angl. *Intracellular Pangenesis*, Chicago : Open Court Publishing Company, 1910.
- de VRIES, H. (1900). Sur la loi de disjonction des hybrides. *Comptes Rendus de l'Académie des Sciences*, 130:845–847.
- DESLOGE, E. (1982). *Classical Mechanics*. Hoboken NJ : John Wiley.
- DIETRICH, M. R. (2000). From gene to genetic hierarchy : Richard Goldschmidt and the problem of the gene. In Beurton *et al.* (2000), pages 91–114.
- DRETSKE, F. (1981). *Knowledge and the Flow of Information*. Oxford : Basil Blackwell.
- DUGAS, R. (1950). *Histoire de la mécanique*. Paris : Dunod.
- DUHEM, P. (1903a). Analyse de l'ouvrage de Ernst Mach : *La mécanique. Étude historique et critique de son développement*. *Bulletin des Sciences Mathématiques*, 27:261–283. Reproduit dans (Duhem, 1903b).
- DUHEM, P. (1903b). *L'évolution de la mécanique*. Introduction et établissement du texte par Anastasios Brenner, avant-propos de Paul Germain, Paris : Vrin, 1992.
- DUHEM, P. (1914). *La Théorie Physique, son objet, sa structure*. Reproduction en fac-similé de la seconde édition (1ère édition : 1906), Paris : Vrin, 1993.

- DUNN, L. C. (1965). *A Short History of Genetics*. New York : McGraw-Hill, rééd. Ames IA : Iowa State University Press, 1991.
- DUPRÉ, J. (1993). *The Disorder of Things*. Cambridge, Mass. : Harvard University Press.
- DYSON, F. J. (1949a). The S matrix in quantum electrodynamics. *Physical Review*, 75:1736–1755.
- DYSON, F. J. (1949b). The radiation theories of Tomonaga, Schwinger, and Feynman. *Physical Review*, 75:486–502.
- DYSON, F. J. (1951). Advanced quantum mechanics. Notes polycopiées d'après des cours donnés à Cornell en 1951.
- ELGIN, C. Z. (1996). *Considerate judgement*. Princeton : Princeton University Press.
- EULER, L. (1744). Methodus inveniendi lineas curvas, maximi minimive proprietate gaudentes sive solutio problematis isoperimetri latissimo sensu accepti. In *Leonhardi Euleri Opera omnia. Seris prima. Opera mathematica*. Volumen Vicesimum quartum. Edidit Constantin Carathéodory. Berne : Orell Füssli Turici, 1952.
- FALK, R. (2004). Applying and extending the notion of genetic linkage. The first fifty years. In Rheinberger et Gaudillière (2004a), pages 34–56.
- FEIGL, H. (1970). The “orthodox” view of theories. In Radner et Winokur (1970), pages 3–16.
- FEYERABEND, P. K. (1962). Explanation, reduction, and empiricism. In FEIGL, H. et MAXWELL, G., éditeurs : *Minnesota Studies in the Philosophy of Science. Scientific Explanation, Space, and Time*, pages 28–97. Minneapolis : University of Minnesota Press.
- FEYERABEND, P. K. (1965a). On the “meaning” of scientific terms. *The Journal of Philosophy*, 62(10):266–274.
- FEYERABEND, P. K. (1965b). Problems of empiricism. In COLODNY, R. G., éditeur : *Beyond the Edge of Certainty*, pages 145–260. Englewood Cliffs, NJ : Prentice-Hall.
- FEYNMAN, R. (1949). Space-time approach to quantum electrodynamics. *Physical Review*, 76(6):769–789.
- FEYNMAN, R. (1965). *The Character of Physical Law*. Cambridge MA : MIT Press.

- FEYNMAN, R., LEIGHTON, R. et SAND, M. (1963). *The Feynman Lectures on Physics*. Addison-Wesley.
- FIELD, H. (1977). Logic, meaning, and conceptual role. *The Journal of Philosophy*, 74(7):379–409.
- FILES, C. (1996). Goodman's rejection of resemblance. *British Journal of Aesthetics*, 36:398–402.
- FINE, A. (1993). Fictionalism. *Midwest Studies in Philosophy*, 18:1–18.
- FLORIDI, L. (2004). Information. In FLORIDI, L., éditeur : *The Blackwell Guide to the Philosophy of Computing and Information*, pages 40–61.
- FLORIDI, L. (2009). Semantic conception of information. In ZALTA, E. N., éditeur : *The Stanford Encyclopedia of Philosophy*. URL : <http://plato.stanford.edu/archives/sum2009/entries/information-semantic/>.
- FODOR, J. et LEPORE, E. (1992). *Holism : A Shoppers' Guide*. Oxford : Blackwell.
- FODOR, J. et LEPORE, E. (2002). *The Compositionality Papers*. Oxford : Clarendon Press.
- FOX KELLER, E. (1983). *A Feeling for the Organism*. New York : Freeman.
- FOX KELLER, E. (1985). *Reflections on Gender and Science*. New Haven : Yale University Press.
- FOX KELLER, E. (1989). *Secrets of Life, Secrets of Death*. New York : Routledge.
- FOX KELLER, E. (1995). *Refiguring Life*. New York : Columbia University Press.
- FRASER, C. (1983). J.L. Lagrange's early contributions to the principles and methods of mechanics. *Archive for History of Exact Sciences*, 28(3):197–241.
- FRASER, C. (1985a). D'Alembert's principle : the original formulation and application in d'Alembert's *Traité de Dynamique* (1743). *Centaurus*. Réimp. in (Fraser, 1997), 28:31–61, 145–159.
- FRASER, C. (1985b). J.L. Lagrange's changing approach to the foundations of the calculus of variations. *Archive for History of Exact Sciences*, 32(2):151–191.
- FRASER, C. (1994). The origins of Euler's variational calculus. *Archive for History of Exact Sciences*. Réimp. in (Fraser, 1997), 47:103–141.

- FRASER, C. (1997). *Calculus and Analytical Mechanics in the Age of the Enlightenment*. Farnham, UK : Ashgate ; Variorum Collected Studies Series.
- FREGE, G. (1894). *Die Grundlagen der Arithmetik*. Breslau : W. Koebner.
- FRIEDMAN, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, 71(1):5–19.
- FRIEDMAN, M. (1999). *Reconsidering Logical Empiricism*. Cambridge : Cambridge University Press.
- FRIGG, R. (2002). Models and representation : Why structures are not enough. *In Measurement in Physics and Economics Project Discussion Paper Series*. DP MEAS 25/02, London School of Economics.
- FRIGG, R. (2006). Scientific representation and the semantic view of theories. *Theoria*, 55:49–65.
- FRIGG, R. (2009). Models and fiction. À paraître dans *Synthese*.
- GALINON, H. (2009). Les termes théoriques, de Carnap à Lewis. *Philonsorbonne*, 4.
- GALLETTO, D. (1991). Lagrange e le origini della meccanica analitica. *Giornale di Fisica*, XXXII(2-3):1–126.
- GALTON, F. (1872). On blood-relationship. *In Proceedings of the Royal Society of London*, volume 20, pages 394–402.
- GALTON, F. (1875). A Theory of Heredity. *Contemporary Review*, 27:80–95.
- GAYON, J. (1992). *Darwin et l'après-Darwin : une histoire de l'hypothèse de sélection dans la théorie de l'évolution*. Paris : Kimé.
- GAYON, J. (1994). De la mesure à l'ordre ; histoire philosophique du concept d'hérédité. *In* PORTE, M., éditeur : *Passion des formes (hommage à René Thom)*, pages 629–645. Paris : ENS Editions Fontenay St Cloud.
- GAYON, J. (1995). Entre force et structure : genèse du concept naturaliste de l'hérédité. *In* WUNENBURGER, J., éditeur : *Le paradigme de la filiation*, pages 61–75. Paris : L'Harmattan.
- GAYON, J. (1996a). Biométrie. *In* TORT, P., éditeur : *Dictionnaire du darwinisme et de l'évolution*, volume 1, pages 318–329. Paris : PUF.
- GAYON, J. (1996b). De la catégorie de style en histoire des sciences. *Alliage*, 26:13–25.

- GAYON, J. (1997). L'émergence du concept biologique d'hérédité au XIX^e siècle. In BERLAND, J.-P., éditeur : *Faut-il créer un privilège sur le vivant?* (Séminaire européen 26-27 septembre 1997, Montpellier).
- GAYON, J. (2000). From measurement to organization : A philosophical scheme for the history of the concept of heredity. In Beurton *et al.* (2000), pages 69–90.
- GENTNER, D. et STEVENS, A., éditeurs (1983). *Mental Models*. Mahwah NJ : Lawrence Erlbaum Associates.
- GIERE, R. N. (1979). *Understanding Scientific Reasoning*. Holt, Rinhart and Winston.
- GIERE, R. N. (1988). *Explaining Science. A Cognitive Approach*. Chicago : The University of Chicago Press.
- GIERE, R. N., éditeur (1992). *Cognitive Models of Science. Minnesota Studies in the Philosophy of Science*. Minneapolis : University of Minnesota Press.
- GIERE, R. N. (2006). *Scientific Perspectivism*. Chicago : The University of Chicago Press.
- GIERE, R. N. (2009). Why scientific models should not be regarded as works of fiction. In Suárez (2009).
- GIORDAN, A., éditeur (1989). *Histoire de la biologie*, volume 2. Paris : Lavoisier.
- GODFREY-SMITH, P. (2006). The strategy of model-based science. *Biology and Philosophy*, 21:725–740.
- GOLDSCHMIDT, R. (1917). Crossing over ohne Chiasmotypie? *Genetics*, 2:82–95.
- GOLDSTEIN, H. (1950/2002). *Classical Mechanics*. Reading, Mass : Addison-Wesley, 3^{ème} édition (en collaboration avec C. P. Poole, Jr et J. L. Safko), 2002.
- GOLDSTINE, H. (1980). *A History of the Calculus of the Variations from the 17th Century through the 19th Century*. New York : Springer-Verlag.
- GOMBRICH, E. H. (1960). *Art and Illusion*. Princeton : Princeton University Press.
- GOMBRICH, E. H. (1982). *The Image and the Eye*. Ithaca NY : Cornell University Press.
- GOODMAN, N. (1968/1976). *Languages of Art. An Approach to a Theory of Symbols*. Hackett Publishing Company (1976, 2nd edition).

- GOPNIK, A. (1996). The scientist as child. *Philosophy of Science*, 63(4):485–514.
- GREENBERG, M. et HARMAN, G. (2006). Conceptual role semantics. In LEPORE, E. et SMITH, B. S., éditeurs : *The Oxford Handbook of Philosophy of Language*, pages 295–322. Oxford : Clarendon Press.
- GRIESEMER, J. R. (1990). Material models in biology. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2:79–93.
- GRIESEMER, J. R. et WIMSATT, W. C. (1989). Picturing Weismannism : A case study of conceptual evolution. In RUSE, M., éditeur : *What the Philosophy of Biology Is. Essays for David Hull*, pages 75–137. Dordrecht : Kluwer Academic Publishers.
- HACKING, I., éditeur (1981). *Scientific Revolutions*. Oxford : Oxford University Press.
- HACKING, I. (1982). Language, truth, and reason. In HOLLIS, M. et LUKES, S., éditeurs : *Rationality and Relativism*, pages 48–66. Cambridge (Mass.) : MIT Press.
- HACKING, I. (1983a). The accumulation of styles of scientific reasoning. In HEINRICH, D., éditeur : *Kant oder Hegel*, pages 453–462. Stuttgart : Klett-Cotta.
- HACKING, I. (1983b). *Representing and Intervening*. Cambridge : Cambridge University Press.
- HACKING, I. (1992a). Statistical language, statistical truth and statistical reason : The self-authentication of a style of reasoning. In MACMULLIN, H., éditeur : *Social Dimensions of Science*, pages 130–157. Notre Dame IN : Notre Dame University Press.
- HACKING, I. (1992b). “style” for historians and philosophers. *Studies in History and Philosophy of Science*, 23:1–20.
- HACKING, I. (1999). *The Social Construction of What?* Harvard University Press, traduction française de B. Jurdant, *Entre science et réalité. La construction sociale de quoi ?*, Paris : La Découverte.
- HACKING, I. (2003). “vrai”, les valeurs et les sciences. In Changeux (2003), pages 201–214.
- HAMILTON, W. R. (1834). On a general method in dynamics. *Philosophical Transactions of the Royal Society*, 2:247–308.

- HANSON, N. R. (1958). *Patterns of Discovery : An Inquiry into the Conceptual Foundations of Science*. Cambridge : Cambridge University Press.
- HANSON, N. R. (1963). *The Concept of the Positron*. New York : Cambridge.
- HARMAN, G. (1982a). Conceptual role semantics. *Notre Dame Journal of Formal Logic*, 23(2):242–256.
- HARMAN, G. (1987). (non-solipsistic) conceptual role semantics. In LEPORE, E., éditeur : *New Directions in Semantics*. Maryland Heights MO : Academic Press.
- HARMAN, P. M. (1982b). *Energy, force and matter : the conceptual development of nineteenth-century physics*. Cambridge : Cambridge University Press.
- HARMAN, P. M., éditeur (1990). *The Scientific Letters and Papers of James Clerk Maxwell*. Cambridge : Cambridge University Press.
- HAUGELAND, J. (1981). Analog and analog. *Philosophical Topics*, 12:213–225.
- HAUGELAND, J. (1991). Representational genera. In RAMSEY, W., STICH, S. et RUMELHART, D., éditeurs : *Philosophy and Connectionist Theory*. Hillsdale : Lawrence Erlbaum. Réimp. dans Haugeland, 1998, pp. 171-206.
- HAUGELAND, J. (1998). *Having Thought*. Cambridge : Harvard University Press.
- HEISENBERG, W. (1942). *Philosophie. Le manuscrit de 1942*. Paris : Éditions du Seuil. Introduction et traduction française de Catherine Chevalley, 1998.
- HEMPEL, C. G. (1950). Problems and changes in the empiricist criterion of meaning. *Revue Internationale de Philosophie*, 4:41–63.
- HEMPEL, C. G. (1951). The concept of cognitive significance : A reconsideration. *Proceedings of the American Academy of Arts and Sciences*, 80:61–77.
- HEMPEL, C. G. (1952). *Fundamentals of Concept Formation in the Empirical Sciences, International Encyclopedia of Unified Science*, volume 2. Chicago : University of Chicago Press.
- HEMPEL, C. G. (1958). The theoretician's dilemma. In FEIGL, H., SCRIVEN, M. et MAXWELL, G., éditeurs : *Concepts, Theories, and the Mind-Body Problem. Minnesota Studies in the Philosophy of Science, vol. 2*. Minneapolis : University of Minnesota Press. Rééd. in Hempel (1965b), pages 173-226.
- HEMPEL, C. G. (1965a). Aspects of scientific explanation. (Hempel, 1965b), pages 331–496.

- HEMPEL, C. G. (1965b). Empiricist criteria of cognitive significance : Problems and changes. In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. Trad. fr. Pierre Jacob in (Jacob, 1980a), pages 101–122. New York : MacMillan.
- HEMPEL, C. G. (1966). *Philosophy of Natural Science*. Prentice Hall. Traduction française par Bertrand Saint-Sernin, Paris : Armand Colin, 1972.
- HEMPEL, C. G. (1970). On the “standard conception” of scientific theories. In Radner et Winokur (1970), pages 142–163.
- HEMPEL, C. G. et OPPENHEIM, P. (1948). Studies in the logic of explanation. *Philosophy of Science*. Reproduit avec une postface dans (Hempel, 1965b), pp. 245–295, 15:135–175.
- HERTZ, H. (1894). *Die Prinzipien der Mechanik. Gesammelte Werke, vol. III*. Leipzig : Barth. Traduction française inédite établie par A. Barberousse et A. Bienvenu. La pagination citée est celle de l’édition allemande.
- HESSE, M. (1966). *Models and Analogies in Science*. Notre Dame IN : Notre Dame University Press.
- HESSE, M. (1982). Comments on Kuhn’s “Commensurability, comparability, communicability”. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2:704–711.
- HILBERT, D. (1899). *Grundlagen der Geometrie*. Leipzig : Teubner.
- HIRSCHFELD, L. A. et GELMAN, S. A., éditeurs (1994). *Mapping the mind : Domain specificity in cognition and culture*. Cambridge : Cambridge University Press.
- HOLLIDAY, R. (1964). A mechanism for gene conversion in fungi. *Genetics Research*, 5:282–303.
- HOPKINS, R. (1998). *Picture, Image and Experience*. Cambridge : Cambridge University Press.
- HUGHES, R. (1989). *The Structure and Interpretation of Quantum Mechanics*. Cambridge MA : Harvard University Press.
- HULL, D. L. (1972). Reduction in genetics - biology of philosophy ? *Philosophy of Science*, 39(4):491–499.

- HULL, D. L. (1974). *The Philosophy of Biological Science*. Englewood Cliffs, NJ : Prentice-Hall.
- HULL, D. L. (1979). Reduction in genetics. *Philosophy of Science*, 46(2):316–320.
- HULL, D. L. (1982). Exemplars and scientific change. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2:479–503.
- HUMPHREYS, Paul, W. (2002). Computational models. *Philosophy of Science*, 69:S1–S11.
- HUMPHREYS, Paul, W. (2004). *Extending Ourselves. Computational Science, Empiricism, and Scientific Method*. Oxford : Oxford University Press.
- HUTCHINS, E. (1995). *Cognition in the Wild*. Cambridge : MIT Press.
- JACOB, F. (1970). *La logique du vivant. Une histoire de l'hérédité*. Paris : Gallimard.
- JACOB, P. (1980a). *De Vienne à Cambridge*. Paris : Gallimard.
- JACOB, P. (1980b). *L'empirisme logique. Ses antécédents, ses critiques*. Paris : Editions de Minuit, coll. Propositions.
- JACOBI, C. G. J. (1837). Note sur l'intégration des équations différentielles de la dynamique. *Comptes Rendus de l'Académie des Sciences de Paris*, 5:61–67.
- JACOBI, C. G. J. (1842). Sur un nouveau principe général de la mécanique analytique. *Comptes rendus de l'Académie des sciences de Paris*, 15:202–205.
- JANSSENS, F. A. (1909). Spermatogénèse dans les batraciens. v. la théorie de la chiasmotypie. nouvelles interprétations les cineses de maturation. *La Cellule*, 25: 387–411.
- JOANNSEN, W. (1909). *Elemente der exakten Erblichkeitslehre*. Iena : Gustav Fischer.
- JOHNS, O. D. (2005). *Analytical Mechanics for Relativity and Quantum Mechanics*. Oxford : Oxford University Press.
- JOHNSON-LAIRD, P. N. (1983). *Mental Models. Towards a cognitive science of language, inference, and consciousness*. Cognitive Science Series. Cambridge MA : Harvard University Press.
- JOHNSON-LAIRD, P. N. (1996). Images, models, and propositional representations. In de VEGA, M., INTONS-PETERSON, M. J., JOHNSON-LAIRD, P. N., DENIS, M. et MARSCHARK, M., éditeurs : *Models of Visuospatial Cognition*, chapitre 3, pages 90–127. Oxford : Oxford University Press.

- KAISER, D. (2000). Stick-figure realism : Conventions, reification, and the persistence of Feynman diagrams, 1948-1964. *Representations*, 70:49–86.
- KAISER, D. (2005). *Drawing Theories Apart : The Dispersion of Feynman Diagrams in Postwar Physics*. Chicago : University of Chicago Press.
- KASS, L. B. et BONNEUIL, C. (2004). Mapping and seeing : Barbara McClintock and the linking of genetics and cytology in maize genetics, 1928-1935. In Rheinberger et Gaudillière (2004a), pages 91–118.
- KATZ, M. (2008). Analog and digital representation. *Minds and Machines*, 18:403–408.
- KIRCHHOFF, G. (1877). *Vorlesungen über mathematische Physik. Mechanik*. Leipzig.
- KIRSH, D. (1991). When is information explicitly represented ? In HANSON, P., éditeur : *Information, Language, and Cognition*. Vancouver : University of British Columbia.
- KITCHER, P. (1978). Theories, theorists, and theoretical change. *The Philosophical Review*, 87(4):519–547.
- KITCHER, P. (1981). Explanatory unification. *Philosophy of Science*, 48:507–531.
- KITCHER, P. (1982a). Genes. *The British Journal for the Philosophy of Science*, 33(4):337–359.
- KITCHER, P. (1982b). Implications of incommensurability. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2:698–703.
- KITCHER, P. (1983). *The Nature of Mathematical Knowledge*. Oxford : Oxford University Press.
- KITCHER, P. (1984). 1953 and all that. A tale of two sciences. *The Philosophical Review*, 93(3):335–373.
- KITCHER, P. (1989). Explanatory unification and causal structure. In Salmon et Kitcher (1989), pages 410–505.
- KITCHER, P. (1993). *The Advancement of Science*. Oxford : Oxford University Press.
- KLEINMUNTZ, D. et SCHKADE, D. (1993). Information displays and decision processes. *Psychological Science*, 4(4):221–227.

- KLINE, M. (1972). *Mathematical Thought from Ancient to Modern Times*. Oxford : Oxford University Press.
- KOHLER, R. E. (1994). *Lords of the Fly : Drosophila Genetics and the Experimental Life*. Chicago : University of Chicago Press.
- KRIPKE, S. (1980). *Naming and Necessity*. Cambridge MA : Harvard University Press.
- KUHN, T. S. (1962/1970c). *The Structure of Scientific Revolutions*. Chicago : University of Chicago Press. Seconde édition augmentée d'une postface, 1970.
- KUHN, T. S. (1969). Second thoughts on paradigms. *In* Suppe (1974/1977b), pages 459–482.
- KUHN, T. S. (1970a). Postface de 1969. *In* Kuhn (1962/1970).
- KUHN, T. S. (1970b). Reflections on my critics. *In* Lakatos et Musgrave (1970), pages 231–278. Rééd. *in* Kuhn (2000), pages 123–175.
- KUHN, T. S. (1976). Theory-change as structure-change : Comments on the Sneed formalism. *Erkenntnis*, 10:179–199. Rééd. *in* Kuhn (2000), pages 176–195.
- KUHN, T. S. (1977a). *The Essential Tension*. Chicago : University of Chicago Press.
- KUHN, T. S. (1977b). Objectivity, value judgment, and theory choice. *In* Kuhn (1977a).
- KUHN, T. S. (1982). Commensurability, comparability, communicability. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2: 669–688.
- KUHN, T. S. (2000). *The Road Since Structure*. Chicago : University of Chicago Press.
- KULVICKI, J. (2003). Image structure. *The Journal of Aesthetics and Art Criticism*, 61(4):323–340.
- KULVICKI, J. (2006a). *On images. Their structure and content*. Oxford : Clarendon Press.
- KULVICKI, J. (2006b). Pictorial representation. *Philosophy Compass*, 1(6):535–546.
- KULVICKI, J. (2009). Knowing with images : Medium and message. À paraître dans *Philosophy of Science*.

- LADYMAN, J. (1998). What is structural realism ? *Studies in History and Philosophy of Science*, 29:409–424.
- LAGRANGE, J.-L. (1760a). Application de la méthode exposée dans le mémoire précédent à la solution de différents problèmes de dynamique. *Miscellanea Taurinensia* t. 2, 1760-1761. Œuvres, vol. 1, pp. 363-468.
- LAGRANGE, J.-L. (1760b). Essai sur une nouvelle méthode pour déterminer les maxima et les minima des formules intégrales indéfinies. *Miscellanea Taurinensia* t. 2, 1760-1761. Œuvres, vol. 1, pp. 333-362.
- LAGRANGE, J.-L. (1764). Recherches sur la libration de la Lune dans lesquelles on tâche de résoudre la question proposée par l'Académie Royale des Sciences, pour le prix de l'année 1764. *Prix de l'Académie Royale des Sciences de Paris tome IX 1780* (1777) Œuvres., 6:5–61.
- LAGRANGE, J.-L. (1788). *Mécanique analytique*. Fac-similé de l'édition originale de 1788 (Paris, chez la Veuve Desaint), Paris : Jacques Gabay, 1989.
- LAKATOS, I. (1970). Falsification and the methodology of scientific research programmes. *In* Lakatos et Musgrave (1970).
- LAKATOS, I. et MUSGRAVE, A., éditeurs (1970). *Criticism and the Growth of Knowledge*. Cambridge : Cambridge University Press.
- LAKOFF, G. et JOHNSON, M. (1980). *Metaphors We Live By*. Chicago University Press.
- LANCZOS, C. (1970). *The variational principles of mechanics*. New York : Dover, 4ème édition. Première édition parue en 1949, Toronto : University of Toronto Press.
- LAPLACE, P.-S. (1843-1846). *Traité de mécanique céleste*. Paris, Imprimerie Royale, édition en *fac simile*, Paris : Jacques Gabay.
- LARKIN, J. H. (1983). The role of problem representation in physics. *In* GENTNER, D. et STEVENS, A. L., éditeurs : *Mental Models*. Mahwah NJ : Lawrence Erlbaum Associates.
- LARKIN, J. H. (1989). Display-based problem solving. *In* KLAHR, D. et KOTOVSKY, K., éditeurs : *Complex Information Processing : The Impact of Herbert A. Simon*. Mahwah NJ : Lawrence Erlbaum Associates.

- LARKIN, J. H., McDERMOTT, J., SIMON, D. et SIMON, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208:1335–1342.
- LATOUR, B. (1995). *La science en action. Introduction à la sociologie des sciences*. Paris : La Découverte.
- LATOUR, B. et WOOLGAR, S. (1979). *Laboratory Life. The Social Construction of Scientific Facts*. Newbury Park CA : Sage Publications.
- LAUDAN, L. (1977). *Progress and Its Problems*. Berkeley CA : University of California Press.
- LAUDAN, L. (1981). A confutation of convergent realism. *Philosophy of Science*, 48:19–49.
- LAUDAN, L., DONOVAN, A., LAUDAN, R., BARKER, P., BROWN, H., LEPLIN, J., THAGARD, P. et WYKSTRA, S. (1986). Scientific change : philosophical models and historical research. *Synthese*, 69(2):141–223.
- LE BIHAN, S. (2006). La conception sémantique des théories scientifiques 1 : En quoi a-t-elle pu se dire “sémantique” ? *Matière Première*, 1.
- LE BIHAN, S. (2008). La conception sémantique des théories scientifiques 2 : Comment peut-elle rester “sémantique” ? *Matière Première*, 3.
- LEE, J. (1999). Words and pictures – goodman revisited. In PATON, R. et NIELSON, I., éditeurs : *Visual Representations and Interpretations*. London : Springer-Verlag.
- LEIBNIZ, G. W. (1765). *Nouveaux essais sur l'entendement humain*. Paris : GF Flammarion, 1990.
- LENAY, C., éditeur (1990). *La découverte des lois de l'hérédité (1862-1900) - Une anthologie*. Paris : Press Pocket.
- LEWIS, D. (1971). Analog and digital. *Noûs*, 5(3):321–327.
- LEWONTIN, R. C. (1992). Genotype and phenotype. In KELLER, E. F. et LLOYD, E. A., éditeurs : *Keywords in Evolutionary Biology*, pages 137–144. Cambridge MA : Harvard University Press.
- LLOYD, E. A. (1988). *The Structure and Confirmation of Evolutionary Theory*. New York : Greenwood Press.
- LOPES, D. (1996). *Understanding Pictures*. Oxford : Oxford University Press.

- LOWENSTAMM, J., éditeur (2003). *Grammaire et gabarits. Recherches linguistiques de Vincennes*.
- LÜTZEN, J. (2003). Between rigor and applications : developments in the concept of function in mathematical analysis. In NYE, M., éditeur : *Cambridge History of Science, vol. 5 : The modern physical and mathematical sciences*, pages 468–487.
- LÜTZEN, J. (2005). *Mechanistic Images in Geometric Form : Heinrich Hertz's Principles of Mechanics*. Oxford : Oxford University Press.
- LYNCH, M. (1988). The externalized retina : Selection and mathematization in the visual documentation of objects in the life sciences. *Human Studies*, 11:201–234.
- LYNCH, M. et WOOLGARD, S., éditeurs (1990). *Representation in Scientific Practice*. Cambridge MA : MIT Press.
- MACH, E. (1883). *Die Mechanik in Ihrer Entwicklung Historisch-Kritisch Dargestellt*. Chicago : Open Court Publishing Company. Traduction anglaise de Thomas J. McCormack, 1893. Sixième édition, 1960.
- MACH, E. (1886). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*. Iena : G. Fischer.
- MACROBERTS, M. H. (1985). Was Mendel's paper on *Pisum* neglected or unknown? *Annals of Science*, 42:339–345.
- MAGNANI, L. et NERSESSIAN, N. J., éditeurs (2002). *Model-Based Reasoning. Science, Technology, Values*. Berlin : Springer.
- MAGNANI, L., NERSESSIAN, N. J. et THAGARD, P., éditeurs (1999). *Model-Based Reasoning in Scientific Discovery*. Berlin : Springer.
- MAIENSCHIN, J. (1991). From presentation to representation in E.B. Wilson's *The Cell*. *Biology and Philosophy*, 6:227–254.
- MARTIN-ROBINE, F. (2006). *Histoire du principe de moindre action : trois siècles de principes variationnels de Fermat à Feynman*. Paris : Vuibert.
- MASTERMAN, M. (1970). The nature of a paradigm. In Lakatos et Musgrave (1970), pages 59–89.
- MAXWELL, J. C. (1855). On Faraday's lines of force. In Harman (1990).
- MAXWELL, J. C. (1890). *The Scientific Papers of James Clerk Maxwell*. Cambridge : Cambridge University Press. Réédition Dover, 1965.

- MAXWELL, J. C. (1920). *Matter and motion*. Dover, notes et appendices de Sir Joseph Larmor.
- MAYR, E. (1982). *The Growth of Biological Thought. Diversity, evolution and inheritance*. Cambridge MA : The Bellknap Press of Harvard University Press.
- MCCCLINTOCK, B. (1930). A cytological demonstration of the location of an interchange between two non-homologous chromosomes of *zea mays*. *Proceedings of the National Academy of Sciences of the United States of America*, 16:791–796.
- MCKINSEY, J., SUGAR, A. et SUPPES, P. (1953). Axiomatic foundations of classical particle mechanics. *Journal of Rational Mechanics and Analysis*, 2(2):253–272.
- MENDEL, G. (1866). Versuche über Pflanzenhybriden. *Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr 1865*, Abhandlungen, pages 3–47. Lu à la conférence de la Société d'Histoire Naturelle de Brünn en février et mars 1865.
- MILLER, G. et JOHNSON-LAIRD, P. N. (1976). *Language and Perception*. MIT : Cambridge MA.
- MONGIN, P. (2003). L'axiomatisation et les théories économiques. *Revue économique*, 54(1):99–138.
- MORANGE, M. (2003). *Histoire de la biologie moléculaire*. Paris : La découverte (première édition : 1994).
- MORGAN, M. S. et MORRISON, M., éditeurs (1999). *Models as Mediators : Perspectives on Natural and Social Science*. Cambridge : Cambridge University Press.
- MORGAN, T. H. (1910a). Chromosomes and heredity. *American Naturalist*, 44:449–296.
- MORGAN, T. H. (1910b). Sex-limited inheritance in *drosophila*. *Science*, 32:1.
- MORGAN, T. H. (1911a). An attempt to analyze the constitution of the chromosomes on the basis of sex-limited inheritance in *drosophila*. *Journal of Experimental Zoology*, 13:79.
- MORGAN, T. H. (1911b). The origin of five mutations in eye color in *drosophila* and their modes of inheritance. *Science*, 33:534.
- MORGAN, T. H. (1911c). Random segregations versus coupling in Mendelian inheritance. *Science*, 34:384.

- MORGAN, T. H. (1917). The theory of the gene. *The American Naturalist*, 51:513–544.
- MORGAN, T. H. (1919). *The Physical Basis of Heredity*. Philadelphia : J.B. Lippincott Company.
- MORGAN, T. H. (1928). *The Theory of the Gene*. New Haven : Yale University Press, second édition.
- MORGAN, T. H. (1935). The relation of genetics to physiology and medicine. *Smithsonian Institution. Annual Report*, pages 345–359.
- MORGAN, T. H. et BRIDGES, C. B. (1916). Sex-linked inheritance in drosophila. *Carnegie Institute Washington Publications*, 237.
- MORGAN, T. H. et CATTELL, E. (1912). Data for the study of sex-limited inheritance in drosophila. *Journal of Experimental Zoology*, 13:79.
- MORGAN, T. H., STURTEVANT, A. H., MULLER, H. J. et BRIDGES, C. B. (1915). *The Mechanism of Mendelian Heredity*. Henry Holt and Company.
- MOULINES, C.-U. (1975). A logical reconstruction of simple equilibrium thermodynamics. *Erkenntnis*, 9.
- MOULINES, C. U. (2002). Introduction : Structuralism as a program for modelling theoretical science. *Synthese*, 130:1–11.
- MOULINES, C.-U. (2006). *La philosophie des sciences. L'invention d'une discipline, fin XIX^e - début XXI^e siècle*. Paris : Rue d'Ulm.
- MOULINES, C.-U. et SNEED, J. (1979). Suppes' philosophy of physics. In *Patrick Suppes*, pages 59–91. Berlin : Springer.
- MULLER, H. J. (1914). A gene for the fourth chromosome of drosophila. *Journal of Experimental Zoology*, 17:325–336.
- MULLER, H. J. (1916). The mechanism of crossing-over. *American Naturalist*, 50(I, II, III, IV):193–221, 284–305, 350–366, 421–434.
- MULLER, H. J. (1920). Are the factors of heredity arranged in a line ? *American Naturalist*, 54:97–121.
- MULLER, H. J. (1951). The development of the gene theory. In DUNN, L. C., éditeur : *Genetics in the 20th Century. Essays on the Progress of Genetics During its First 50 Years*, pages 77–99. New York : MacMillan.

- MULLIGAN, J. F. (1998). The reception of Heinrich Hertz's *Principles of Mechanics* by his contemporaries. In BAIRD, D., HUGHES, R. et NORDMANN, A., éditeurs : *Heinrich Hertz : Classical Physicist, Modern Philosopher*. Dordrecht : Kluwer Academic Publishers.
- NAGEL, E. (1961). *The Structure of Science : Problems in the Logic of Scientific Explanation*. Harcourt, Brace and World, Inc.
- NERSESSIAN, N. (2008). *Creating Scientific Concepts*. Cambridge MA : MIT Press.
- NERSESSIAN, N. J. (1984). *Faraday to Einstein : Constructing Meaning in Scientific Theories*. Dordrecht : Kluwer.
- NERSESSIAN, N. J. (1992a). How do scientists think? capturing the dynamics of conceptual change in science. In GIERE, R. N., éditeur : *Cognitive Models of Science*, pages 3–45. Minneapolis : University of Minnesota Press.
- NERSESSIAN, N. J. (1992b). In the theoretician's laboratory : Thought experimenting as mental modeling. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2:291–301.
- NERSESSIAN, N. J. (2002a). Kuhn, conceptual change, and cognitive science. In NICHOLS, T., éditeur : *Thomas Kuhn, Contemporary Philosophers in Focus*, pages 178–211. Cambridge : Cambridge University Press.
- NERSESSIAN, N. J. (2002b). Maxwell and “the method of physical analogy” : Model-based reasoning, generic abstraction and conceptual change. In MALAMENT, D. B., éditeur : *Essays in the History and Philosophy of Science and Mathematics*, pages 129–166. Chicago : Open Court Publishing Company.
- NERSESSIAN, N. J. (2007). Mental modeling in conceptual change. In VOSNIADOU, S., éditeur : *The Handbook of Conceptual Change*. Mahwah NJ : Lawrence Erlbaum Associates.
- NEURATH, O. (1933). Protokollsätze. *Erkenntnis*, 3:204–214. Traduction française « Énoncés protocolaires », in (Soulez, 1985).
- NEWTON, I. (1687). *Philosophiae Naturalis Principia Mathematica*. Traduction française par la Marquise du Châtelet, 1759. Fac similé aux éditions Jacques Gabay, 1990.
- NORTH, J. (2009a). Structure in classical mechanics. Manuscrit.

- NORTH, J. (2009b). The structure of physics. À paraître dans *Journal of Philosophy*.
- OLBY, R. (1985). *Origins of Mendelism*. Chicago : University of Chicago Press, second édition.
- OLBY, R. (1987). William Bateson's introduction of Mendelism to England : a reassessment. *British Journal for the History of Science*, 20(67):399–420.
- OUELBANI, M. (2006). *Le Cercle de Vienne*. Paris : PUF.
- PAINTER, T. P. (1934a). The morphology of the X chromosome in salivary glands of *Drosophila melanogaster* and a new type of chromosome maps for this element. *Genetics*, 19:448–469.
- PAINTER, T. P. (1934b). A new method for the study of chromosomes aberrations and the plotting of chromosome maps in *Drosophila melanogaster*. *Genetics*, 19:175–188.
- PAINTER, T. P. (1935). The morphology of the third chromosome in the salivary gland of *Drosophila melanogaster* and a new cytological map for this element. *Genetics*, 20:301–326.
- PANZA, M. (1990). The evolution of Lagrange's research programme in the analytical foundations of mechanics with particular respect to its realisation in the "théorie des fonctions analytiques". *Historia Scientiarum*, 1-2 (1991) et 1-3 (1992).
- PANZA, M. (2003). The origins of analytical mechanics in 18th century. In JAHNKE, H., éditeur : *A History of Analysis*. American Mathematical Society and London Mathematical Society, s. 1, pages 131–153.
- PASCAL, B. (1670). *Pensées*. Paris : Classiques Garnier, texte de l'édition Brunschvicg, 1958.
- PEACOCKE, C. (1986). The inaugural address : Analogue content. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 70:1–17.
- PEACOCKE, C. (1987). Depiction. *The Philosophical Review*, 96(3):383–410.
- PEACOCKE, C. (1992). *A Theory of Concepts*. Cambridge MA : MIT Press.
- PEARSON, K. (1898). Mathematical contributions to the theory of evolution. on the law of ancestral heredity. *Proceedings of the Royal Society of London*, 62:386–412.
- PEGNY, M. (2005). Le principe de moindre action de Maupertuis à Lagrange. Mémoire de D.E.A., Université Paris 10 - Nanterre. Mémoire de maîtrise.

- PERINI, L. (2004). Convention, resemblance and isomorphism : understanding visual representations. In MALCOLM, G., éditeur : *Studies in Multidisciplinarity*, volume 2, pages 37–47.
- PERINI, L. (2005). Explanation in two dimensions : diagrams and biological explanation. *Biology and Philosophy*, 20:257–269.
- PICKERING, A. (1992). *Science as Practice and Culture*. Chicago : Chicago University Press.
- POINCARÉ, H. (1902). *La Science et l'Hypothèse*. Paris : Flammarion, collection « Champs », 2000.
- POINCARÉ, H. (1905). *La Valeur de la Science*. Paris : Flammarion, collection « Champs », 2000.
- POISSON, S. D. (1838). *Traité de Mécanique*. Bruxelles : Société belge de librairie, Hauman et compagnie.
- POLANYI, M. (1958). *Personal Knowledge : Towards a Post-Critical Philosophy*. Chicago : University of Chicago Press.
- POTTER, H. et DRESSLER, D. (1979). Branch migration in recombination. *Cold Spring Harbor Symposium of Quantitative Biology*, 43:957–990.
- PRESTON, J. (2008). Mach and Hertz's mechanics. *Studies in History and Philosophy of Science*, 39(1):91–101.
- PSILLOS, S. (1999). *Scientific Realism. How Science Tracks Truth*. London : Routledge.
- PULTE, H. (1989). *Das Prinzip der kleinsten Wirkung und die Kraftkonzeptionen der rationalen Mechanik. Eine Untersuchung zur Grundlegungsproblematik bei Leonhard Euler, Pierre Louis Moreau de Maupertuis und Joseph-Louis Lagrange*. Studia Leibnitiana. Sonderheft 19. Stuttgart : Franz Steiner Verlag.
- PUNNETT, R. C. (1913). Reduplication series in sweet peas. *Journal of Genetics*, 3:77–103.
- PUTNAM, H. (1962). What theories are not. In NAGEL, E., SUPPES, P. et TARSKI, A., éditeurs : *Logic, Methodology and Philosophy of Science : Proceedings of the 1960 International Congress*, pages 240–251. Palo Alto CA : Stanford University Press.

- PUTNAM, H. (1973). Explanation and reference. In PEARCE, G. et MAYNARD, P., éditeurs : *Conceptual Change*. Dordrecht : D. Reidel Publishing Company.
- PUTNAM, H. (1975). The meaning of “meaning”. In *Language, Mind, and Knowledge : Minnesota Studies in the Philosophy of Science VII*.
- PUTNAM, H. (1981). *Reason, Truth, and History*. Cambridge : Cambridge University Press.
- QUINE, W. V. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60:20–43.
- QUINE, W. V. (1953a). *From a Logical Point of View*. Cambridge MA : Harvard University Press.
- QUINE, W. V. (1953b). On what there is. In Quine (1953a).
- QUINE, W. V. (1969). Epistemology naturalized. In *Ontological Relativity and Other Essays*. New York : Columbia University Press.
- QUINE, W. V. (1975). On empirically equivalent systems of the world. *Erkenntnis*, 9.
- RADNER, M. et WINOKUR, S., éditeurs (1970). *Analyses of Theories and Methods of Physics and Psychology. Minnesota Studies in the Philosophy of Science, vol. 4*. Minneapolis : University of Minnesota Press.
- RAIMOND, J.-M. (2000). Électromagnétisme et relativité. II. Mécanique analytique. Notes de cours disponibles sur <http://www.phys.ens.fr/cours/notes-de-cours/jmr/electromagnetisme.htm>.
- RAMSEY, F. (1929). Theories. In BRAITHWAITE, R. B., éditeur : *The Foundations of Mathematics and Other Logical Essays*, pages 212–236. Patterson, NJ : Littlefield and Adams, 1950.
- RAPOPORT, A. (1958). Various meanings of “theory”. *American Political Science Review*, 52:927–988.
- REICHENBACH, H. (1924). *Axiomatik der relativischen Raum-Zeit Lehre*. Braunschweig : F. Vieweg.
- REICHENBACH, H. (1928). *Philosophie der Raum-Zeit-Lehre*. Berlin : W. de Gruyter.
- RESNIK, M. D. (1997). *Mathematics as a Science of Patterns*. Oxford : Oxford University Press.

- RHEINBERGER, H.-J. (2000). Gene concepts : Fragments from the perspective of molecular biology. *In* Beurton *et al.* (2000), pages 219–239.
- RHEINBERGER, H.-J. et GAUDILLIÈRE, J.-P., éditeurs (2004a). *Classical Genetics Research and Its Legacy. The mapping cultures of twentieth-century genetics*. London : Routledge.
- RHEINBERGER, H.-J. et GAUDILLIÈRE, J.-P., éditeurs (2004b). *From Molecular Genetics to Genomics. The mapping cultures of twentieth-century genetics*. London : Routledge.
- RHEINBERGER, H.-J. et MÜLLER-WILLE, S. (2007). *Gene*. The Stanford Encyclopedia of Philosophy.
- ROSENBERG, A. (1985). *The Structure of Biological Science*. Cambridge/New York : Cambridge University Press.
- ROSENBERG, A. (1994). *Instrumental Biology or the Disunity of Science*. Chicago : University of Chicago Press.
- ROSENBERG, A. (2000). *Darwinism in Philosophy, Social Science and Policy*. Cambridge : Cambridge University Press.
- RUMELHARD, G. (1986). *La génétique et ses représentations dans l'enseignement*. Berne. Francfort-sur-Main. New York : Peter Lang.
- RUSSELL, B. et WHITEHEAD, A. N. (1910/1913). *Principia Mathematica*. Cambridge : Cambridge University Press.
- SALMON, W. (1989). Four decades of scientific explanation. *In* Salmon et Kitcher (1989), pages 3–220.
- SALMON, W. et KITCHER, P., éditeurs (1989). *Minnesota Studies in the Philosophy of Science, Vol. 13 : Scientific Explanation*. Minneapolis : University of Minnesota Press.
- SCHAFFNER, K. F. (1969a). Correspondence rules. *Philosophy of Science*, 36(3):280–290.
- SCHAFFNER, K. F. (1969b). The Watson-Crick model and reductionism. *The British Journal for the Philosophy of Science*, 20(4):325–348.
- SCHIER, F. (1986). *Deeper into Pictures*. Cambridge : Cambridge University Press.

- SCHILPP, P. A., éditeur (1963). *The Philosophy of Rudolf Carnap*. Library of Living Philosophers, vol. 11. Chicago : Open Court Publishing Company.
- SCHWINGER, J. (1948a). On quantum-electrodynamics and the magnetic moment of the electron. *Physical Review*, 73:416–417.
- SCHWINGER, J. (1948b). Quantum electrodynamics. *Physical Review*, 74:439–461.
- SHAPER, D. (1964). The structure of scientific revolutions. *The Philosophical Review*, 73:383–394.
- SHAPER, D. (1966). Meaning and scientific change. In COLODNY, R. G., éditeur : *Mind and Cosmos : Essays in Contemporary Science and Philosophy*, pages 41–85. Pittsburgh PA : University of Pittsburgh Press.
- SHAPIRO, S. et SCHAFER, S. (1985). *Léviathan and the Air-Pump, Hobbes, Boyle, and the Experimental Life*. Princeton NJ : Princeton University Press.
- SHAPIRO, S. (1983). Mathematics and reality. *Philosophy of Science*, 50(4):523–548.
- SIMON, H. A. et LARKIN, J. H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11:65–99.
- SNEED, J. (1971). *The Logical Structure of Mathematical Physics*. Dordrecht : D. Reidel Publishing Company. 2ème édition 1979.
- SNEED, J. (1975). *The Logical Structure of Mathematics*. London : Routledge and Kegan Paul.
- SNEED, J. (1976). Philosophical problems in the empirical science of science : a formal approach. *Erkenntnis*, 10.
- SOKAL, A. (1996). Transgressing the boundaries : Towards a transformative hermeneutics of quantum gravity. *Social Text*, 46-47:217–252.
- SOULEZ, A., éditeur (1985). *Manifeste du Cercle de Vienne et autres écrits*. Paris : PUF.
- SPELKE, E. S. (1991). Physical knowledge in infancy : Reflections on Piaget's theory. In CAREY, S. et GELMAN, R., éditeurs : *The Epigenesis of Mind : Essays on Biology and Cognition*. Mahwah NJ : Lawrence Erlbaum Associates.
- STEGMÜLLER, W. (1976). *The Structure and Dynamics of Theories*. New York : Springer-Verlag.

- STERN, C. (1931). Zytologisch-genetische Untersuchungen als beweis für die morgansche Theorie des Faktorenaustausches. *Biologische Zentralblatte*, 51:547–587.
- STRICKBERGER, M. (1985). *Genetics*. New York : Macmillan.
- STURTEVANT, A. H. (1913a). The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology*, 14:43–59.
- STURTEVANT, A. H. (1913b). A third group of linked gene in *drosophila ampelaphila*. *Science*, 37:990–992.
- STURTEVANT, A. H. (1914). The reduplication hypothesis as applied to *Drosophila*. *American Naturalist*, 48:535–549.
- STURTEVANT, A. H. (1915). The behavior of the chromosomes as studied through linkage. *Zeitschrift für inductive Abstammungs- und Vererbungslehre*, 13:234–287.
- STURTEVANT, A. H. (1917). Crossing over without Chiasmatype? *Genetics*, 2:301–304.
- STURTEVANT, A. H. (1965). *A History of Genetics*. New York : Harper & Row.
- STURTEVANT, A. H., BRIDGES, C. B. et MORGAN, T. H. (1919). The spatial relations of genes. *Proceedings of the National Academy of Sciences*, 5:168–173.
- SUÁREZ, M. (1999). Theories, models, and representation. In MAGNANI, L., NERSESIAN, N. et THAGARD, P., éditeurs : *Model-Based Reasoning in Scientific Discovery*, pages 75–83. London : Kluwer Academic/Plenum Publishers.
- SUÁREZ, M. (2003). Scientific representation : against similarity and isomorphism. *International Studies in the Philosophy of Science*, 17(3):225–244.
- SUÁREZ, M. (2004). An inferential conception of scientific representation. *Philosophy of Science*, 71 (5), Proceedings of the 2002 Biennial Meetings of the Philosophy of Science Association:767–779.
- SUÁREZ, M., éditeur (2009). *Fictions in Science. Philosophical Essays on Modeling and Idealization*. London : Routledge.
- SUPPE, F. (1971). On partial interpretation. *The Journal of Philosophy*, 68(3):57–76.
- SUPPE, F. (1974/1977b). *The Structure of Scientific Theories*. Champaign IL : University of Illinois Press. Seconde édition augmentée d’une introduction critique et d’une postface, 1977.

- SUPPE, F. (1977a). Introduction. *In* Suppe (1974/1977b), pages 3–232.
- SUPPE, F. (1989). *The Semantic View of Theories and Scientific Realism*. Champaign IL : University of Illinois Press.
- SUPPES, P. (1957). *Introduction to Logic*. Princeton : Van Nostrand.
- SUPPES, P. (1959). Axioms for relativistic kinematics with or without parity. *In* HENKIN, L., SUPPES, P. et TARSKI, A., éditeurs : *The Axiomatic Method with Special Reference to Geometry and Physics*, pages 191–307.
- SUPPES, P. (1960). A comparison of the meaning and use of models in mathematics and the empirical sciences. *Synthese*.
- SUPPES, P. (1962). Models of data. *In* NAGEL, E., SUPPES, P. et TARSKI, A., éditeurs : *Logic, Methodology and Philosophy of Science : Proceedings of the 1960 International Congress*, pages 252–261. Palo Alto CA : Stanford University Press.
- SUPPES, P. (1967). What is a scientific theory ? *In* MORGENBESSER, S., éditeur : *Philosophy of Science Today*, pages 55–67. New York : Basic Books.
- SUPPES, P. (1968). The desirability of formalization in science. *The Journal of Philosophy*, 65(20):651–664.
- SUPPES, P. (1988). Representation theory and the analysis of science. *Philosophia Naturalis*, 25:254–268.
- SUPPES, P. (2002). *Representation and Invariance of Scientific Structures*. Stanford : CSLI Publications.
- SUZUKI, D. T., GRIFFITHS, A. J., MILLER, J. H., LEWONTIN, R. C. et GELBART, W. M. (1986). *Introduction to Genetic Analysis*. New York : W.H. Freeman and Company.
- TARSKI, A. (1935). Le concept de vérité dans les langages formalisés. *In* tarski1972. Traduction française dirigée par G. Granger, Paris : Armand Colin.
- TARSKI, A. (1936a). La construction d’une sémantique scientifique. *In* tarski1972. Traduction française dirigée par G. Granger, Paris : Armand Colin.
- TARSKI, A. (1936b). Sur le concept de conséquence logique. *In* tarski1972. Traduction française dirigée par G. Granger, Paris : Armand Colin.
- THAGARD, P. (1988). *Computational Philosophy of Science*. Cambridge MA : MIT Press.

- THAGARD, P. (1992). *Conceptual Revolutions*. Princeton : Princeton University Press.
- THOMPSON, P. (1989). *The Structure of Biological Theories*. New York : State University of New York.
- THOMPSON, P. (2007). Formalisations of evolutionary biology. In MATTHEN, M. et STEPHENS, C., éditeurs : *Handbook of the Philosophy of Science. Philosophy of Biology.*, volume 2. Philosophy of Biology de *Handbook of the Philosophy of Science*. Elsevier.
- THOMSON, W. (1884). *Lectures on molecular Dynamics, and the Wave-Theory of Light*. Baltimore MD : John Hopkins University.
- THOMSON, W. et TAIT, P. G. (1888). *Treatise on Natural Philosophy*. Cambridge : Cambridge University Press.
- THOMSON-JONES, M. (2006). Models and the semantic view. *Philosophy of Science*, 73(5):524–535.
- THREADGOLD, L. (1976). *The Ultrastructure of the Animal Cell*. Oxford : Pergamon Press.
- TOULMIN, S. (1959). Criticism in the history of science : Newton on absolute space, time, and motion. *Philosophical Review*, 68:1–29, 203–207.
- TOULMIN, S. (1961). *Foresight and Understanding : An Enquiry into the Aims of Science*. Bloomington : Indiana University Press.
- TUFTE, E. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT : Graphics press.
- TUFTE, E. (1990). *Envisioning Information*. Cheshire, CT : Graphics press.
- TVERSKY, B. (2001). Spatial schemas in depictions. In GATTIS, M., éditeur : *Spatial schemas and abstract thought*, pages 79–111. Cambridge MA : MIT Press.
- TVERSKY, B. (2004). Semantics, syntax, and pragmatics of graphics. In HOLMQVIST, K. et ERICSSON, Y., éditeurs : *Language and visualisation*, pages 141–158. Lund : Lund University Press.
- UFFINK, J. (2004). Boltzmann's work in statistical physics. In ZALTA, E. N., éditeur : *The Stanford Encyclopedia of Philosophy*.

- VAIHINGER, H. (1911). *The philosophy of "as if"*. London : Kegan Paul. Traduction anglaise 1924.
- van FRAASSEN, B. C. (1980). *The Scientific Image*. Oxford : Oxford University Press.
- van FRAASSEN, B. C. (1987). The semantic approach to scientific theories. In NER-SESSIAN, N. J., éditeur : *The Process of Science*, pages 105–124. Martinus Nihoff.
- van FRAASSEN, B. C. (1989). *Laws and Symmetry*. Oxford : Clarendon Press.
- van FRAASSEN, B. C. (1991). *Quantum Mechanics : An Empiricist View*. Oxford : Oxford University Press.
- van FRAASSEN, B. C. (2008). *Scientific Representation : Paradoxes of Perspective*. Oxford : Oxford University Press.
- VARIGNON, P. (1700a). Du mouvement en général par toutes sortes de courbes ; et des forces centrales, tant centrifuges que centripètes, nécessaires aux corps qui les décrivent. *Histoire de l'Académie Royale des Sciences avec les Mémoires de mathématique et de physique pour la même année, année 1700*, pages 83–101.
- VARIGNON, P. (1700b). Manière générale de déterminer les forces, les vitesses, les espaces et les temps, une seule de ces quatre choses étant donnée dans toutes sortes de mouvements rectilignes variés à discrétion. *Histoire de l'Académie Royale des Sciences avec les Mémoires de mathématique et de physique pour la même année, année 1700*, pages 22–27.
- VISSER, H. (1999). Boltzmann and Wittgenstein, or How pictures became linguistic. *Synthese*, 119:135–156.
- von HELMHOLTZ, H. L. (1847). *Über die Erhaltung der Kraft*. Berlin.
- von NEUMANN, J. et BURKS, A. W. (1966). *Theory of Self-Reproducing Automata*. Champaign IL : University of Illinois Press.
- von TSCHERMAK, E. (1900). Über künstliche Kreuzung bei *Pisum sativum*. *Zeitschrift für das landwirtschaftliche Versuchswesen in Österreich*, 3:465–555.
- WAGNER, P. (2002a). Carnap et la logique de la science. In Wagner (2002b), pages 246–298.
- WAGNER, P., éditeur (2002b). *Les philosophes et la science*. Paris : Gallimard.

- WALTON, K. (1990). *Mimesis as make-believe : On the foundations of the representational arts*. Cambridge MA : Harvard University Press.
- WATERS, C. K. (1990). Why the anti-reductionist consensus won't survive : The case of classical Mendelian genetics. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1:125–139.
- WATERS, C. K. (1994). Genes made molecular. *Philosophy of Science*, 61(2):163–185.
- WATERS, C. K. (Spring 2007). Molecular genetics. In ZALTA, E. N., éditeur : *The Stanford Encyclopedia of Philosophy*.
- WATSON, J. (1968). *The Double Helix*. New York : Atheneum.
- WATSON, J. et CRICK, F. (1953a). Genetical implications of the structure of deoxyribose nucleic acid. *Nature*, 171.
- WATSON, J. et CRICK, F. (1953b). Molecular structure of nucleic acids : a structure for deoxyribose nucleic acid. *Nature*, 171.
- WEINBERG, S. (1996). Sokal's hoax. *New York Review of Books*, 8:11–15.
- WEINSTEIN, A. (1962). The reception of Mendel's paper by his contemporaries. In *Proceedings of the Tenth International Congress of the History of Science*, pages 997–1001. Ithaca, Paris, 1964.
- WEISMANN, A. (1889). *Essays Upon Heredity and Kindred Biological Problems*. Oxford : Clarendon Press.
- WESFALL, R. (1971). *Force in Newton's Physics : the Science of Dynamics in the Seventeenth Century*. Londres.
- WHITE, B. Y. (1983). Sources of difficulty in understanding Newtonian dynamics. *Cognitive Science*, 7:41–65.
- WHITESIDE, D. (1970). *The Mathematical Principles Underlying Newton's "Principia Mathematica"*. Glasgow : University of Glasgow.
- WHITTAKER, E. (1959). *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*. Cambridge : Cambridge University Press. 4ème édition.
- WHORF, B. (1956). *Language, Thought, and Reality : Selected Writings of Benjamin Lee Whorf*. Cambridge MA : MIT Press.

- WILSON, A. (1989). Hertz, Boltzmann and Wittgenstein reconsidered. *Studies in History and Philosophy of Science*, 20(245-263).
- WILSON, E. (1925). *The Cell in Development and Heredity*. New York : MacMillan.
- WIMSATT, W. C. (1976). Reductionism, levels of organization and the mind-body problem. In GLOBUS, G., MAXWELL, G. et SAVODNIK, I., éditeurs : *Consciousness and the Brain : A Scientific and Philosophical Inquiry*, pages 199–267. New York : Plenum Press.
- WIMSATT, W. C. (1987). False models as means to truer theories. In NITECKI, M. et HOFFMAN, A., éditeurs : *Neutral Models in Biology*, pages 23–55. Oxford : Oxford University Press.
- WIMSATT, W. C. (1990). Taming the dimensions - visualizations in science. *PSA : Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2: 111–135.
- WIMSATT, W. C. (1992). Golden generalities and co-opted anomalies : Haldane vs. Muller and the drosophila group on the theory and practice of linkage mapping. In *The Founders of Evolutionary Genetics. A centenary reappraisal*, pages 101–166. Dordrecht : Kluwer Academic Publishers.
- WITTGENSTEIN, L. (1922). *Tractatus Logico-Philosophicus*. London : Routledge and Kegan Paul.
- WITTGENSTEIN, L. (1953). *Philosophische Untersuchungen*. Frankfurt : Suhrkamp.
- WOLLHEIM, R. (1974). *On Art and the Mind*. Cambridge MA : Harvard University Press.
- WOLLHEIM, R. (1980). *Art and Its Objects*. Cambridge : Cambridge University Press, 2nde éd.
- WOLLHEIM, R. (1993). *The Mind and Its Depths*. Cambridge MA : Harvard University Press.
- WOLLHEIM, R. (1998). On pictorial representation. *The Journal of Aesthetics and Art Criticism*, 56:217–226.
- WOODHOUSE, N. (1987). *Introduction to Analytical Dynamics*. Oxford : Oxford University Press.

- WOODY, A. (2004). More telltale signs : What attention to representation reveals about scientific explanation. *Philosophy of Science*, 71 (5), Proceedings of the 2002 Biennial Meetings of the Philosophy of Science Association:780–793.
- WRIGHT, S. (1932). The role of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics*, pages 355–366.
- YOURGRAU, W. et MANDELSTAM, S. (1968). *Variational Principles in Dynamics and Quantum Theory*. Philadelphia, PA : Saunders.
- YOUSCHKEVITCH, A. (1976). The concept of function up to the middle of the nineteenth century. *Archive for the History of Exact Sciences*, 16:37–85.
- ZHANG, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, 21(2):179–217.
- ZUKAV, G. (1979). *The Dancing Wu Li Masters : An Overview of the New Physics*. William Morrow and Company.

Index des noms propres

- Achinstein, P., 198, 209, 211, 330
 Adams, E.W., 247
 Allen, G., 434, 436, 447, 485, 499, 512
 Anderson, J.R., 398
 Archimède, 138
 Arnold, V.I., 37, 77

 Bütschli, O., 445
 Baigrie, B., 345
 Balzer, W., 240, 269, 270, 272, 336
 Banks, E.C., 125
 Barberousse, A., 30, 32, 97, 106, 117, 157, 322, 335
 Barroso Filho, W., 106
 Bateson, W., 429, 432, 440, 447, 482–485, 487, 499, 503, 504, 511, 513–515, 525, 530, 534–536
 Beatty, J., 240
 Beckner, M., 240
 Bell, J., 148
 Bernouilli, J., 54, 59
 Beth, E., 239
 Blachowicz, J., 358
 Black, M., 212, 281, 316
 Blackmore, J.T., 125
 Blay, M., 28, 103, 104
 Block, N., 313, 358
 Bloor, D., 13
 Bogdan, R., 241
 Bohr, N., 171, 216
 Boltzmann, L., 122, 123, 134–136, 169–174, 216, 217, 294, 318
 Bonnet, C., 193
 Bonneuil, C., 505
 Boscovitch, R., 38, 145, 241
 Bottazini, U., 37
 Bourbaki (groupe), 242
 Bouveresse, J., 193
 Boveri, T., 444–446
 Brandom, R., 98
 Brannigan, A., 429
 Brenner, A., 144
 Bridges, C.B., 485, 499, 501–504, 516, 518, 521, 532, 533
 Bridgman, P., 200
 Budd, M., 356
 Burian, R., 426, 458
 Butterfield, J., 36, 37, 52, 54, 56, 57, 59, 61, 77, 80, 81, 124, 148, 150, 172

 Callender, C., 369
 Callon, M., 13
 Campbell, N.R., 200, 212
 Cannon, W.A., 446
 Cantor, G., 193
 Carey, S., 13, 313
 Carlson, E.A., 483, 485, 486, 491, 499–501
 Carlyle, T., 2
 Carnap, R., 4, 192–211, 218–221, 225–227, 229, 230, 233, 237, 245
 Cartwright, N., 11, 43, 48, 281, 316, 322
 Casati, R., 380

- Castle, W.E., 433, 511, 514–536
 Cat, J., 96
 Cattell, E., 487, 494, 496, 499
 Chakravartty, A., 254
 Chapuis-Schmitz, D., 193, 201, 219
 Chevalley, C., 171, 216
 Clark, A., 99
 Cock, A., 484
 Cohen, I.B., 103
 Cohen, J., 360, 369
 Cohen, R., 125
 Coleman, W., 484
 Comte, C., 106
 Correns, C., 429
 Creighton, H.B., 504–507
 Crick, F., 325, 382, 426
 Crombie, A.C., 158

 D'Agostino, S., 134
 d'Alembert, J., 28, 54, 109
 Dahan-Dalmedico, A., 106, 107, 112, 113, 152
 Dalton, J., 294
 Darden, L., 429, 434, 443, 444, 458, 481, 484, 501, 526, 527
 Darrigol, O., 51, 54, 55, 79, 82, 110–112, 115, 117
 Darwin, C., 433, 434, 437, 438
 David, P., 431, 432, 435, 443, 458, 471, 526, 530, 534, 539
 Davidson, D., 311
 de Courtenay, N., 123
 De Gandt, F., 103
 de Maupertuis, P.L.M., 106, 113, 157
 de Regt, H.W., 134, 170, 173
 de Vries, H., 429, 433, 437, 446
 Delbrück, M., 477
 Descartes, R., 241, 293
 Desloge, E., 148
 Dieks, D., 173
 Dietrich, M.R., 512
 Dressler, D., 473
 Dretske, F., 358, 360
 Dugas, R., 103
 Duhem, P., 3–4, 6, 123, 133, 135, 143–147, 152, 155, 163–167, 169, 216, 218, 222, 233
 Dunn, L.C., 429, 434, 445, 485, 487, 499, 501
 Dupre, J., 458
 Dyson, F.J., 395, 408–411, 416, 417, 532
 Einstein, A., 24, 89, 90, 301–305, 307, 309, 310, 350, 351, 359
 Elgin, C.Z., 322
 Euclide, 191
 Euler, L., 28, 38, 58, 59, 106–108, 110, 111, 113, 115, 116, 145, 150
 Falk, R., 501
 Feigl, H., 238
 Feyerabend, P.K., 283, 302, 305, 311
 Feynman, R., 101, 114, 115, 130, 167–169, 349, 392–396, 406, 408–413, 415–417, 532
 Fick, R., 445
 Field, H., 313
 Files, C., 356
 Fine, A., 231, 322
 Fischer, R.A., 339
 Fleming, W., 445
 Floridi, L., 360
 Fodor, J., 313
 Fox Keller, E., 96
 Fraser, C., 106, 110
 Frege, G., 193
 Friedman, M., 173

- Frigg, R., 249, 258, 281, 316, 319–326
- Galilée, 138
- Galinon, H., 206, 207
- Galletto, D., 106
- Galton, F., 434, 435, 437
- Gaudillière, J.-P., 480
- Gayon, J., 158, 427, 429, 433–438, 474, 475, 538
- Gelman, S.A., 317
- Gentner, D., 317
- Giere, R.N., 43, 45, 48, 238, 240, 244, 253–255, 259, 316, 317, 319, 322, 326, 345
- Giordan, A., 430, 434
- Godfrey-Smith, P., 281, 319–323
- Goldschmidt, R., 504, 511–514, 525
- Goldstein, H., 32, 36, 50, 56, 63, 67, 68, 148
- Goldstine, H., 106
- Gombrich, E.H., 356
- Goodman, N., 183, 348, 350–356, 358–360, 363, 364, 366, 367, 370, 372, 373, 376, 385, 401, 417
- Gopnik, A., 13, 313
- Greenberg, M., 313, 345
- Griesemer, J.R., 281, 345, 346, 480
- Hacking, I., 12, 158, 283, 316, 550
- Haldane, J.B.S., 339, 503
- Hamilton, W.R., 27, 28, 30, 33, 58, 86, 87, 100, 103, 117–119, 412, 415
- Hanson, N.R., 173, 283, 284
- Harman, G., 313, 345
- Haugeland, J., 358, 366
- Heisenberg, W., 171
- Hempel, C.G., 4, 187, 195, 197–199, 201, 202, 204–206, 214, 215, 234, 237, 238, 255, 550
- Hertwig, O., 445
- Hertz, H., 3, 33, 92, 99, 100, 122–124, 133–146, 152, 154, 155, 160, 163, 169, 170
- Hesse, M., 212, 281, 316
- Hilbert, D., 191, 193, 202, 242
- Hirschfeld, L.A., 317
- Holliday, R., 472
- Hooke, R., 45
- Hopkins, R., 356
- Hughes, R., 240, 263
- Hull, D.L., 346, 428, 458
- Humphreys, P.W., 7, 12, 74, 81, 166, 241, 246, 281, 282, 328, 336–343, 345, 536
- Hutchins, E., 99
- Jacob, F., 427, 439, 450, 452, 458, 538
- Jacob, P., 193, 194, 196, 206
- Jacobi, C.G.J., 28, 58, 59
- Janssens, F.A., 447, 451, 487, 488, 512
- Joannsen, W., 429, 435
- Johns, O.D., 148
- Johnson, M., 96, 383
- Johnson-Laird, P.N., 313, 405
- Kaiser, D., 14, 15, 392, 393, 395, 408–411
- Kant, E., 241
- Kass, L.B., 505
- Katz, M., 358
- Kirchhoff, G., 122
- Kirsh, D., 365
- Kitcher, P., 4, 12, 18, 214, 311, 426, 428, 456, 457, 459–476, 541, 542, 545
- Kleinmutz, D., 390
- Kline, M., 37
- Kohler, R.E., 434, 480, 485, 499, 500
- Kripke, S., 311

- Kuhn, T.S., 11, 13, 17, 29, 44, 81, 89, 90, 119, 167, 182, 219, 281–306, 308–311, 315, 316, 318, 323, 326–337, 345, 415, 416, 460, 475, 538, 545, 551
 Kulvicki, J., 353, 354, 356, 378, 389, 399, 400
 Lützen, J., 37, 134
 Ladyman, J., 37
 Lagrange, J.L., 28, 30, 33, 54, 58, 59, 85–87, 94, 99, 100, 103–120, 122, 125, 126, 130, 131, 133, 137–139, 144–148, 150, 154–158, 168
 Lakatos, I., 283, 284, 336
 Lakoff, G., 96, 383
 Lanczos, C., 36, 49, 52, 54–56, 58–60, 67, 77, 147–150
 Laplace, P.S., 28, 118, 144, 146, 158
 Larkin, J.H., 98, 387, 390, 398
 Latour, B., 13, 545
 Laudan, L., 78, 283, 284
 Lawrence, Sir T., 361
 Le Bihan, S., 187, 239, 257, 259, 266
 Lee, J., 371
 Legendre, A.-M., 67
 Leibniz, G.W., 103, 104, 133, 145, 146, 148
 Lenay, C., 433, 434
 LePore, E., 313
 Lewis, D., 358
 Lewontin, R.C., 435, 442
 Lloyd, E.A., 240, 244, 246, 263
 Lopes, D., 356
 Louvrier de Lajolais, A., 364
 Lowenstamm, J., 337
 Ludwig, P., 322
 Lynch, M., 345, 375, 385, 386
 Mach, E., 27, 30, 32, 100, 122–124, 126–134, 140, 142
 Macroberts, M.H., 429
 Magnani, L., 212, 317
 Maienschein, J., 452
 Mandelstam, S., 117
 Martin-Robine, F., 106
 Masterman, M., 284
 Maull, N., 458
 Maxwell, J.C., 96, 122, 172, 213, 216
 Mayr, E., 434, 499
 McClintock, B., 504–507
 McClung, C.E., 447
 McKinsey, J., 242
 McKinsey, J., 269
 Mendel, G., 426, 428–433, 436, 438–445, 448, 450, 474
 Miller, G., 313
 Mongin, P., 192
 Montgomery, T., 445
 Morange, M., 426, 435, 439, 458, 477
 Morgan, M.S., 43, 48, 316, 338
 Morgan, T.H., 426, 427, 429, 433, 435, 436, 439, 443, 446–449, 453–455, 480–482, 485–492, 497–501, 503, 504, 510–518, 521, 523–528, 531–536
 Morrison, M., 43, 48, 316, 338
 Moulines, C.U., 240, 246, 269, 270, 272, 336
 Muller, H., 434, 477, 485, 503, 523–526
 Mulligan, J.F., 142
 Naegeli, C., 444
 Nagel, E., 202, 203, 237, 209 – –342
 Nersessian, N., 13, 96, 212, 214, 317
 Neurath, O., 194, 197
 Newton, I., 7, 24, 25, 27, 28, 37, 38, 87, 99,

- 103–105, 115–118, 128–130, 132, 133, 137–139, 144–148, 155, 241
- North, J., 36, 48, 64, 74, 75, 77–79, 124
- Olby, R., 429, 432, 434
- Ouelbani, M., 193
- Painter, T.P., 505, 507–511, 528, 529, 531
- Panza, M., 106
- Pascal, B., 163
- Peacocke, C., 313, 356, 358
- Pearson, K., 437
- Pegny, M., 106, 108, 110, 113
- Perini, L., 345, 356, 370, 371
- Pickering, A., 13
- Poincaré, H., 2, 122, 127, 128, 135, 136
- Poisson, S.D., 28, 118, 144, 146, 338
- Polanyi, M., 290
- Potter, H., 473
- Preston, J., 134
- Psillos, S., 206
- Pulte, H., 106, 115, 118–120, 131, 156
- Punnett, R.C., 432, 450, 462, 482–485, 487, 511, 530
- Putnam, H., 187, 194, 198, 209, 237, 311
- Quine, W.V., 36, 195, 197, 215, 226, 267, 298, 304
- Raimond, J.-M., 36, 61, 65, 148
- Ramsey, F., 206, 208, 229
- Rapoport, A., 257
- Reichenbach, H., 200
- Resnik, M.D., 37
- Rheinberger, H.-J., 480
- Rosenberg, A., 195, 215, 458
- Rumelhard, G., 426
- Russell, B., 193
- Salmon, W., 4, 214
- Schaffer, S., 13
- Schaffner, K.F., 234, 428, 458
- Schier, F., 356
- Schilpp, P.A., 193
- Schkade, D., 390
- Schlick, M., 193
- Schulze, M., 445
- Schwinger, J., 395, 410, 417
- Seeger, R., 125
- Shapere, D., 278, 283, 305–308, 314, 475
- Shapin, S., 13
- Shapiro, S., 37
- Simon, H.A., 98, 387, 390, 398
- Sneed, J., 240, 246, 269, 270, 272, 336
- Sokal, A., 158
- Soulez, A., 193
- Spelke, E.S., 13
- Stegmüller, W., 336
- Stern, C., 505
- Stevens, A., 317
- Stevens, N.M., 447
- Strasburger, E., 445
- Strickberger, M., 463, 465
- Sturtevant, A., 433, 434, 444, 480, 481, 485, 491–499, 501, 503, 512, 513, 515, 516, 518, 523, 526, 532, 533, 535
- Suárez, M., 249, 258, 316, 319, 322
- Sugar, A., 242, 269
- Suppe, F., 188, 246, 251, 264–267, 284, 287, 330, 331
- Suppes, P., 188, 239–251, 257, 258, 260–262, 264, 269, 274, 334, 336, 339
- Sutton, W., 444–446
- Suzuki, D.T., 473
- Tait, P.G., 122, 139
- Tarski, A., 196, 240, 242

- Thagard, P., 13, 212, 317
Thompson, P., 240, 246, 263
Thomson, W. (Lord Kelvin), 117, 122, 123, 139, 172
Thomson-Jones, M., 249, 258
Threadgold, L., 375
Toulmin, S., 283
Tufte, E., 385
Tversky, B., 383, 401
Uffink, J., 123
Vaihinger, H., 231, 322
van Fraassen, B.C., 188, 190, 240, 244, 246, 248, 250, 251, 253–256, 258, 263–267, 274
Varignon, P., 28, 103–105
Varzi, A., 380
Vilmorin, L. de, 435
Vilmorin, P.-P. de, 435
Visser, H., 134
von Helmholtz, H.L., 117, 122
von Neumann, J., 173, 338, 340
von Tschermak, E., 429
Waddington, C., 435
Wagner, P., 193
Walton, K., 322
Waters, C.K., 426, 428, 442, 458, 472, 473
Watson, J., 325, 426, 458
Weinberg, S., 158, 550
Weinstein, A., 429
Weismann, A., 346, 433–435, 444
Wellington, A.W. duc de, 361, 363–364
Wesfall, R., 103
White, B.Y., 98
Whitehead, A.N., 193
Whiteside, D., 103
Whittaker, E., 60
Whorf, B., 298
Wilson, A., 134
Wilson, E., 444, 452–454, 485, 501
Wimsatt, W.C., 281, 343, 345, 458, 480, 503, 521, 523, 526, 536
Wittgenstein, L., 193, 195
Wollheim, R., 356
Woodhouse, N., 148
Woody, A., 345, 415
Woolgar, S., 14, 345, 545
Wright, S., 339
Yourgrau, W., 117
Youschkevitch, A., 37
Zhang, J., 390
Zukav, G., 174

Table des figures

1	Système masse-ressort	45
2	Graphe de l'évolution dynamique du système masse ressort	46
3	Schéma du pendule simple	47
4	Schéma de deux pendules liés	50
5	Graphe des trajectoires réelle et variée de l'état d'un système	61
6	Faisceau tangent d'un cercle	76
7	Espace des phases du pendule simple	77
8	Table des matières de la <i>Mécanique analytique</i> de Lagrange (1788). . .	121
9	L'espace à sept points	190
10	Schéma de la théorie scientifique selon l'empirisme logique	238
11	Schéma du pendule simple	320
12	La représentation scientifique selon R. Frigg	321
13	Watson, Crick, et la double hélice	325
14	L'électrocardiogramme normal	350
15	Albert Einstein	350
16	Mémorial d'Einstein à Washington	351
17	Portrait du duc de Wellington	361
18	Graphe de la température, des précipitations et de l'ensoleillement à Paris	362
19	Caricature du duc de Wellington	364
20	Caryotype	368
21	Diagramme moléculaire	370
22	Micrographe électronique de cellules pulmonaires	373
23	Micrographe électronique du protozoaire <i>Giardia</i>	374
24	Mitochondrie : micrographe électronique et schéma	375
25	Schéma de vulgarisation : cellules reproductrices	377

26	Cartes de France	378
27	Schémas de vulgarisation : molécules d'ADN	380
28	Hexagone	380
29	Dogme central de la biologie moléculaire	382
30	Carte des habitudes de recherche des scientifiques sur Internet	384
31	Diagramme de Mach	385
32	Résultats d'un relevé de température : liste et tableaux	389
33	Diagramme de Feynman	394
34	Diagramme de Feynman	394
35	Tableau de températures	400
36	Diagramme de la première loi de Mendel	440
37	Diagramme de la deuxième loi de Mendel	441
38	Entrecroisement de chromosomes chez le batracien	447
39	Modèle de « collier de perles »	449
40	Spermatogénèse chez le batracien	451
41	Schéma de la méiose	452
42	Schéma de la phase réductionnelle de la méiose	454
43	Diagramme de la ségrégation et de l'assortiment indépendant	454
44	Diagramme (arbre) de la première loi de Mendel	463
45	Diagramme de pedigree : caractère dominant	463
46	Diagramme de pedigree : caractère récessif	463
47	Diagramme du pedigree de l'hémophilie dans la famille de la reine Victoria	464
48	Diagramme de pedigree : caractère lié au sexe	464
49	Diagramme de pedigree : deux caractères liés au sexe	465
50	Modèle de Holliday de la recombinaison génétique	473
51	Schéma de l'hypothèse de reduplication	484
52	Drosophiles (photographie)	485
53	Diagramme présentant les résultats du croisement d'un mâle drosophile aux yeux blancs et d'une femelle aux yeux rouges	486
54	Entrecroisements de chromosomes pendant la méiose	488
55	Modèle du collier de « perles »	489
56	Tableau des proportions de crossing-over sur le chromosome <i>X</i> de <i>Drosophila</i>	493
57	Carte génétique du chromosome <i>X</i> de <i>Drosophila</i>	493

58	Tableau présentant les exceptions à l'additivité des fréquences de re-combinaison	494
59	Schéma de l'hypothèse du double crossing-over	494
60	Résultats numériques indiquant des crossing-over doubles	495
61	Diagramme du test back-cross	500
62	Carte génétique des quatres chromosomes de <i>Drosophila</i>	502
63	Graphe des fréquences attendues et observées de doubles crossing-over	503
64	Schémas des chromosomes étudiés par Creighton et McClintock (1931)	506
65	Carte cytologique et génétique de McClintock (1930)	507
66	Diagramme du croisement opéré par Creighton et McClintock (1931). .	507
67	Résultats numériques des expériences de croisement de Creighton et McClintock (1931).	507
68	Chromosome de la glande salivaire de <i>Drosophila</i>	508
69	Chromosome de la glande salivaire de <i>Drosophila</i>	509
70	Carte cyto-génétique du troisième chromosome de <i>Drosophila</i>	510
71	Carte du chromosome <i>X</i> de Morgan et Bridges (1916)	519
72	Modèle du « piège à rat » de Castle (1919)	520
73	Carte cytogénétique du chromosome <i>X</i> de <i>Drosophila</i>	529

Liste des tableaux

1	Le phénomène de liaison partielle	490
2	Positions respectives de Goldschmidt, Bateson et Castle	525

Théories, modes d'emploi. Une perspective cognitive sur l'activité théorique dans les sciences empiriques

Résumé : Qu'est-ce qu'une théorie scientifique, et quelle pertinence cette notion a-t-elle pour l'étude de la connaissance scientifique ? Cette thèse vise à montrer que, si l'on considère les théories comme des outils de représentation et d'inférence, l'analyse de leur contenu doit prendre en compte la manière dont elles sont comprises, en pratique, par leurs utilisateurs, c'est-à-dire par les agents, profanes et experts, qui les apprennent, les développent et les appliquent. Dans une telle perspective, la forme sous laquelle les théories sont présentées, ainsi que le contexte de leur utilisation, deviennent primordiaux. En critiquant certaines approches classiques de la notion de théorie, cette thèse propose des outils d'analyse et définit une méthode pour étudier l'activité théorique ; deux études de cas – en mécanique classique et en génétique – la mettent en œuvre et en montrent la fécondité.

Theories, A User's Guide. A cognitive approach to theorizing in the empirical sciences

Summary : What is a scientific theory ? And how relevant is this notion for a study of scientific knowledge ? This thesis argues that an analysis of the content of theories – conceived of as tools for representing and drawing inferences – should pay attention to the way agents use and understand them, through learning and practice. From this perspective, the form in which theories are displayed, and the context in which they are used, are crucial. This thesis criticizes some classical approaches to the notion of theory, and defines tools of analysis for a study of theorizing ; this method is applied to two case studies – in classical mechanics and in genetics –, which show its fruitfulness.

Discipline : Philosophie

Mots-clés : Philosophie des sciences ; Théorie scientifique ; Histoire des sciences ; Modèle ; Schéma ; Image ; Paradigme ; Cognition scientifique ; Mécanique classique ; Génétique.

Équipe d'accueil : Institut d'Histoire et de Philosophie des Sciences
et des Techniques (UMR 8590), 13 rue du Four, 75006 Paris.

École doctorale : École doctorale de Philosophie de l'Université Paris 1 (ED 280),
1 rue d'Ulm, 75005 Paris.